THE FORMATION AND LOGIC OF QUANTUM MECHANICS

Vol. I

The Formation of Atomic Models
This page is intentionally left blank
THE
FORMATION AND LOGIC
OF
QUANTUM MECHANICS

Vol. I
The Formation of Atomic Models

MITUO TAKETANI

Translated from the Japanese and with Explanatory Notes by Masayuki Nagasaki

World Scientific
New Jersey • London • Singapore • Hong Kong
Preface to the English Edition

This book is the English edition of *The Formation and Logic of Quantum Mechanics* written in Japanese, which consists of the following three volumes:

Vol. I — *The Formation of Atomic Models* — by Taketani, with *Explanatory Notes* by Nagasaki added to its republished edition,

Vol. II — *The Way to Quantum Mechanics* — by Taketani and Nagasaki,


This book was planned by Taketani soon after the World War II, and Vol. I was published in 1948 by a Japanese publishing company, which became unfortunately insolvent afterward. The republication of Vol. I and the publication of the volumes to be connected were not accomplished for a long time in spite of the hope of many readers, because of various circumstances, of which some descriptions are given in the mentioned explanatory notes. Volume I was re-published in 1972 by means of photocomposition, to which *Explanatory Notes* by Nagasaki were added at Taketani’s request, mainly for the sake of students and readers born after the World War II. Volumes II and III were accomplished by the collaboration of Taketani and Nagasaki and published respectively in 1991 and 1993.

The aim of this book is to analyze through what intricate logical process the quantum theory was developed, and to elucidate by what logic quantum mechanics thus established is governed. The method of our analysis is based
on the three stage theory of scientific cognition, which was presented by Take- 
tani to solve the ineffectiveness of the Machist view of science in treating the 
difficulties met with in nuclear physics at that time. According to the Machist 
view, scientific cognition is to know some mathematical laws by means of the 
ystematization of given sensuous experience. The three stage theory gives, on 
the contrary, the view that scientific cognition proceeds through the series of 
coiling turns of the three stages, namely, the phenomenological, substantialistic 
and essentialistic stages.

The latter view was presented in 1934 by Taketani a little before the pro­ 
posal by H. Yukawa of the meson theory, and was developed in a close con­ 
nection with the works by H. Yukawa, S. Sakata and Taketani of constructing 
the meson theory. H. Yukawa describes this situation in his preface with the 
title “The way I proceeded” to the book written by S. Takauchi in Japanese 
as follows: “In the innermost recesses of my heart there was existing for a long 
time a great doubt. It started growing in my heart a little after my graduation 
in 1929. It was, as is described in the chapter with the title “Tenki” (Turning­ 
point) of my autobiography “Tabibito” (Traveler), a fundamental doubt of, or 
a feeling of dissatisfaction with, the quantum theory of field. Briefly expressed, 
it was the consciousness of an essentialistic problem. Afterward, shelving this 
problem for a while, I proceeded to the meson theory which was full of sub­ 
stantialistic nature”.

About the three stage theory, the reader will find more explanations in 
Concluding Remarks of Vol. I and in Explanatory Notes added to Vol. I. Some 
descriptions of the three stage theory are also given here and there in the text 
of the present book, in relation to our concrete analyses of the process of the 
formation of quantum mechanics, and in connection with our elucidation of 
the logic of quantum mechanics. In the preface to each volume mentions are 
made of the outline of our way of analyzing the problems in the epoch treated 
in the volume concerned.

The three stage theory is in sharp contrast to the Machist theory in at­ 
taching importance to the stereo-structural nature of the logic comprehended 
in scientific cognition, instead of the plane-projected view, so to say, of sci­ 
entific cognition taken in the latter theory. From the analyses given in the 
present book, the reader will see how it is important, in discussing quantum 
mechanics as well as its formation, to distinguish the stereo-structural logic 
from the plane-projected view based on the formal logic.
The three stage theory was somewhat introduced internationally, by means of the publication of the English translation of some related original papers in Supplement No. 50 (1971), Progress of Theoretical Physics (Kyoto). A part of the mentioned publication in English was published in German as "Wissenschaftliche Taschenbüchen Bd. 221" (Akademie-Verlag, Berlin, 1980). L. M. Brown and H. Rechenberg give also a certain introduction to the three stage theory in their recent book "The Origin of the Concept of Nuclear Forces" (Institute of Physics Publishing, Bristol and Philadelphia, 1996).

The present English edition of our book has been prepared by Nagasaki as the translator. The quotations from cited papers written in German and French are translated also into English by Nagasaki. His thanks are due to Miss T. Ohsuga of Laboratory of Theoretical Physics, Rikkyo University for her assistance in typewriting.

Mituo Taketani
Masayuki Nagasaki
Dec. 24, 1997

Postcript added in proof: I, M. Nagasaki, would like to thank Prof. M. Konuma of Musashi Institute of Technology for his interest in the present book, and for his kindness to introduce it to Prof. K. K. Phua of World Scientific Publishing Co. Pte. Ltd. I am grateful to Prof. Phua for his goodwill to publish it from the mentioned publishing company. My thanks are also due to Dr. L. Y. Wong as Scientific Editor of the same company for his good offices in its publication.

I add by myself the present postscript, because, to my great sorrow, Prof. Taketani passed away on April 22, 2000. Prof. Taketani was of the same mind as me to express our sincere thanks of Prof. Konuma, Prof. Phua and Dr. Wong in our Preface to the English Edition at the time of reading the first Proof.

Masayuki Nagasaki
September 24, 2000
This page is intentionally left blank
Atomic physics is in every respect one of the greatest achievements of human beings, and is recognized as the one to give human thought a great revolution and a firm foundation. Nevertheless, there has been little work of making clear its logical structure to put the outcome in the property of thought. Natural scientists, who are generally not versed in logical expression and elucidation of high level, have not been able to clarify the significance of the works they did, so that they understand it superficially and their interpretations are quite separated from their works. Philosophers on the other hand have not analyzed the logical structure of physics itself, and have selected fragments of one or another physicist’s words, using those irresponsibly for the sake of their own theories.

Therefore, such a work must be done first of all as to trace the correct process of formation of the atomic theory, and thereby to draw its logic. This is however surely a hard task. I hope that my present book will serve as a footing for this aim.


Volume I treats the process up to the establishment of Rutherford’s model of the atom, in which our Nagaoka made an important work, so great a contribution to world physical society made by Japanese physics in its young days that we can never forget it. Nevertheless, its actual fact is generally unknown.

Also in foreign countries, almost no detailed book on the history of the formation of atomic models before the time of Rutherford has not been seen,
because this subject was regarded as settled already. However, the period mentioned is very important when it is seen from the viewpoint of epistemology. Every epistemological confusion arises from having no appreciation of this period.

In the present book particular attention is paid to reveal clearly on what assumption each paper bases, and what conclusion is deduced by means of what reasoning as well as from what fact, for the purpose of clarifying step by step the process of development in it.

The method of carrying out this purpose has been described in detail in my book "Problems in Dialectic". I shall be glad if the reader refers to this book.

I intend that descriptions shall represent the features of each original paper as far as possible, and also that words shall be faithful to the original papers. In this respect the present book may lack plainness and unity such as those found in textbooks. As to this point, I would like to recommend the reader to consult some textbook of quantum mechanics, such as that by S. Tomonaga, for example, for supplementary information. The knowledge which I suppose to be necessary to understand the present book, is that which is treated in textbooks of general physics for undergraduate students, and that of the Maxwell theory of electromagnetics.

As a work on the history of science, the present one makes only the fundamental study as an initial step. It is still necessary to make clear the relation of science to society and the interaction of science with general philosophical thought. I plan for a study on "the experimental foundation of quantum mechanics" as to the former, and a study on "the role played by idealism and materialism in the formation of atomic physics" as to the latter.

The book "The History of Quantum Mechanics", a great work by the late Mr. Kiyoshi Amano who was my respected senior, has recently been published to my delight. I hope that the reader will read through it once, since into it plenty of descriptions of social background and episodes are woven into, and it is plain to be readable by general readers. I would like to thank Mr. Amano in this occasion for his having given me much knowledge.

As the present book is the first step in my study I am afraid that it may have insufficient points. I would like to get suggestions from those who know well about circumstances of those times.

The present book is written in a specialized way with the use of mathematical equations. But they are put in for the sake of contributing to more
understanding. Those who are not familiar with mathematical equations may well skip the mathematical parts to get the outline of reasoning.

I would like to acknowledge Mr. S. Tomonaga, Mr. S. Sakata and other my seniors and co-workers for their good wishes and advice. I am indebted very much to Mr. S. Nakamura for his help in getting the literature. My thanks are due to Mr. G. Tominga, Miss Yuminami and Miss Ghi for their help about the literature.

I would like to thank Mr. Y. Kuyoshi of the Ginza Pub. Co. for his goodwill to the present book of unpopular subject and for his good offices in its publication. I am grateful to Mr. M. Minakami and Mr. M. Watanabe of the same company for their acceptance of my personal circumstances to offer convenience. I intended to write out the present book in August of last year, but I spent almost one year more in finishing my work. I would like to express my sincere thanks to the workers of the printing office for the trouble they have taken for me.

The author
May 8, 1948
This page is intentionally left blank
Contents

Preface to the English Edition v
Preface ix

Volume I. The Formation of Atomic Models 1

Chapter 1. Quantum of Radiation 3
1. Thermodynamical Investigation of Black Body Radiation 3
   §1. The prehistory of thermal radiation 3
   §2. Kirchhoff’s law 4
   §3. Growth of atomistic viewpoint 6
   §4. Stefan-Boltzmann’s law 6
   §5. Wien’s displacement law 8
2. Atomistic Investigations of Black Body Radiation 12
   §1. Application of atomism 12
   §2. Wien’s distribution law 14
   §3. Planck’s effort 16
   §4. Foundation by Planck of Wien’s distribution formula 17
   §5. Rayleigh’s critique based on the theory of vibration 21
   §6. Breakdown of the classical theory 23
   §7. Planck’s new distribution formula 24
   §8. Planck’s effort to found the new formula 26
   §9. Discovery of energy quantum 27
3. Einstein’s Light Quantum 33
§1. Photoelectric effect .............................................. 33
§2. Einstein’s critique of the classical theory ......................... 34
§3. Derivation of the radiation formula from the classical theory . 35
§4. Einstein’s consideration of Planck’s theory ........................ 37
§5. Two limits of Planck’s formula .................................... 38
§6. Meaning of Wien’s formula ......................................... 38
§7. Treatment of radiation by means of the kinetic theory of gases 41
§8. Stokes’ law and the light quantum ................................ 43
§9. Photoelectric effect and the light quantum ........... 44
§10. Ionization of molecule ........................................ 47
§11. Einstein’s analysis of Planck’s theory ............................ 48

4. The Light Quantum and the Theory of Relativity ................. 50

Chapter 2. The Formation of Atomic Models .......................... 55
1. On Line Spectra .................................................. 55
§1. Experiments on line spectra ...................................... 55
§2. Mathematical law of line spectra ................................. 60
§3. Attempts of deriving spectral laws from dynamical models .... 66

2. Ether Model of the Atom ........................................... 73
§1. Atomism .......................................................... 73
§2. Vortex atoms ..................................................... 75

3. Clarification by the Theory of Electrons .......................... 77
§1. Lorentz’s theory of electrons .................................... 77
§2. The Zeeman effect .............................................. 85
§3. Explanation by Lorentz’s theory of electrons ................... 88
§4. Experimental verification by Zeeman of Lorentz’s theory ... 91
§5. Measurement of $e/m$ ............................................ 93
§6. Larmor’s theory .................................................. 94

4. Model of the Atom without Nucleus ................................ 100
§1. Study of cathode rays by J. J. Thomson ......................... 100
§2. Thomson’s atomic model in terms of corpuscles .............. 106
§3. On the electric charge of electron ................................ 110
§4. Kelvin’s model of the atom ..................................... 115
§5. Permanence and magnetism of atoms ............................ 127
§6. The Thomson model of the atom ................................ 138

5. Model of the Atom with Nucleus .................................. 162
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>§1. Lenard's Dynamiden atom</td>
<td>162</td>
</tr>
<tr>
<td>§2. Nagaoka’s Saturn-like atom</td>
<td>168</td>
</tr>
<tr>
<td>§3. Arguments against Schott’s critique</td>
<td>179</td>
</tr>
<tr>
<td>6. Examinations of Both Models and the Determination of Atomic Model</td>
<td>182</td>
</tr>
<tr>
<td>§1. About the number of electrons in an atom-I</td>
<td>182</td>
</tr>
<tr>
<td>§2. Origin of atomic mass</td>
<td>195</td>
</tr>
<tr>
<td>§3. Investigations of oscillations in atoms and of intensities of spectral lines</td>
<td>197</td>
</tr>
<tr>
<td>§4. The determination by Rutherford</td>
<td>207</td>
</tr>
<tr>
<td>§5. About the number of electrons in an atom-II</td>
<td>221</td>
</tr>
<tr>
<td>Concluding Remarks</td>
<td>231</td>
</tr>
<tr>
<td>Explanatory Notes</td>
<td>234</td>
</tr>
</tbody>
</table>
Volume I

The Formation of Atomic Models
This page is intentionally left blank
Chapter 1

Quantum of Radiation

1. Thermodynamical Investigation of Black Body Radiation\(^1\)

\(\S 1.\ \text{The prehistory of thermal radiation}\)

Let us first take investigations of the emission or absorption of thermal energy by a body at a certain temperature by means of radiation into or from space, namely, investigations of the fact, for example, that the hotter a solid body is heated, the more intensive the heat rays emitted by it becomes, and at a temperature slightly above 500° C (that is at 525° C) it begins to emit red rays of light which are perceptible by eyes, the emitted rays becoming whiter and whiter at higher temperatures. About the investigation of radiations of such nature, there was early Newton’s well-known law of cooling (1701), and there were found afterward Lambert’s law (1760) of the relation of the flow of light to the plane passed through by it, and Prevost’s law (1792) of heat exchange, which says that every one of bodies at one and the same temperature exchanging heat among them by means of radiation, emits an amount of radiation equal to that of radiation it receives from the other bodies.

In 19th century the theory of heat developed fast, and the knowledge of thermal radiation advanced rapidly along with the establishment of the wave theory of light, so that the identity of the ray of heat with the ray of light was

\(^1\)On this subject there has already been published an excellent book by Kiyoshi Amano "The Origin of the Theory of Thermal Radiation and Quantum Theory" (Dai-Nippon Pub. Co., 1943; in Japanese), which the reader is recommended to refer to. In the present section, we shall thus be concerned mainly in the question of logical development. Cf. also Takuzo Sakai "Thermal Radiation" (Iwanami Lecture Series in Physics, Vol. VI B, 1939; in Japanese), in which the matter under consideration is described in a historical way.
made clear. There appeared the discovery of infrared rays by Herschel (1800), Dulong–Petit’s law of cooling (1817), Ritchie’s experiment on emissive power and absorptive power (1833), Ampère’s consideration of a displacement law of thermal radiation. In thermodynamics, Carnot presented the Carnot cycle of heat engine in 1814, and Mayer the conservation law of energy in 1942. In the next year 1943 Joule made clear the mechanical equivalence of heat, and then in 1947 Helmholtz declared the conservation law of energy. The second law of thermodynamics was proposed by Clausius in 1850 and by W. Thomson in 1951, and the kinetic theory of gases was shown by Krönig in 1856 and by Clausius in 1857.

In this way the basis for the study of thermal radiation was nearly established, and from about the middle of the 19th century the knowledge of it was rapidly deepened, to bring about steady progress in the theoretical research on it. From nearly the same time the center of research moved into Germany. This was directly connected with the rapid development of German industry at that time, as is described in detail in Amano’s book cited above. That is, after the Prussia-France War (1870–71) as the turning point, the weight of importance of German economy moved from agriculture to industry, and metallurgical industry, gas industry, illumination industry, etc. demanded and promoted the study of problems of high temperature and thermal radiation.

§2. Kirchhoff’s law

It is Kirchhoff’s law (1859) that makes the starting point in the theory of thermal radiation. A similar law was found by Stewart in 1856 and published in 1861, but it did not play an important role in the development of the study made mainly in the Continent. Kirchhoff established in 1859 with Bunsen the method of spectral analysis. In this study it was made clear that line spectra are characteristics of the respective elements, but he proceeded in the same year, starting from the inversion of Fraunhofer lines, to the direction of the problem of the law of emission and absorption of radiation, that is, the problem of the theory of thermal radiation, of which the basis for theoretical consideration was already provided by thermodynamics, instead of going to the direction of the problem of the cause of line spectra.2)

2) The problem of line spectra is one of the main themes in our later chapters. The subject making the central part of both the problems was thus given its starting point by Kirchhoff in that year.
Kirchhoff developed his theory on the basis of a strict thermodynamical thought experiment to establish the concept of "black body", and made clear the law of activity of emission and that of absorption. The concept of black body played a central role in the later development of the theory of thermal radiation.

Kirchhoff's law states that the ratio of the emissive power $e(\lambda, T)$ to the absorptive power $a(\lambda, T)$ of any body which is in the thermal equilibrium at a temperature, is irrespective of the kind of the matter making up the body, being a function only of the wavelength $\lambda$ of thermal radiation and the temperature $T$. A black body is then defined as an ideal body which completely absorbs every radiation falling on it, so that the emissive power of any black body $e_0(\lambda, T)$ is a definite function only of $\lambda$ and $T$ regardless of the nature of every material thing making it. The ratio of the emissive power to the absorptive power mentioned above of a body in general, is therefore identical with the emissive coefficient of a black body $e_0(\lambda, T)$, because the absorptive power of any black body $a_0(\lambda, T)$ is equal to unity. In other words, one has

$$a_0(\lambda, T) = 1,$$  \hspace{1cm} (1a)

$$e(\lambda, T)/a(\lambda, T) = \text{a universal function} = e_0(\lambda, T).$$  \hspace{1cm} (1b)

From these properties the nature of the radiation which is in thermal equilibrium with a black body at a temperature is determined. Such radiation is called black body radiation.

Kirchhoff's second law states that, when any body is in thermal equilibrium with thermal radiation in vacuum, the thermal radiation is identical, of whatever kind the body is, with black body radiation.

Thus, the radiation field in a cavity enclosed with a wall which absorbs completely but does not emit radiation, is concluded to become black body radiation if there is in the cavity an absorptive body, however small its quantity is. This black body radiation corresponds to the state of maximum entropy. It was left to later studies to determine the universal function $F(\lambda, T)$ representing $e_0(\lambda, T)$, which is dependent only on $\lambda$ and $T$ but is independent of the kind of matter, that is

$$e_0(\lambda, T) = F(\lambda, T).$$  \hspace{1cm} (2)

In the proof of the first law use was made of Kirchhoff's reciprocal theorem. It states that the amount of black body radiation, which emerges from a surface
taken in the field of black body radiation and arrives at another such surface, is equal to the amount of black body radiation, which emerges from the latter surface and arrives at the former surface. The violation of this theorem means a contradiction with the second law of thermodynamics.

Kirchhoff's law should therefore be said to be the application of the second law of thermodynamics to thermal radiation.

§3. Growth of atomistic viewpoint

The foundation of the theory of thermal radiation was thus established. Then, in 1860 the law of distribution of velocity of gas molecules was given by Maxwell, and in 1864 the basic equations of electromagnetic field were proposed by him. At nearly the same time experimental studies of thermal radiation were advanced, and in 1864 the dependence of thermal radiation on temperature was studied experimentally by Tyndall about platinum. In 1865 the concept of entropy was founded by Clausius, so that the theoretical basis of thermodynamics was established. In 1871 the electromagnetic theory of light was proposed by Maxwell. In the same year Bartoli treated the problem of light pressure, discussing its relation to the second law of thermodynamics. In 1877 there appeared Boltzmann’s famous foundation of the second law of thermodynamics by means of statistical mechanics, and the firm foundation was thus given to atomism, that aimed to explain phenomena from the substantialistic structure of matter in view of materialistic molecules, contrary to the phenomenological tendency from the time of Kirchhoff, that is, energetics that rested on thermodynamical bases. Hard controversies were developed between the two separate schools of energetics and atomism in opposition to each other.3)

Through these controversies, there grew up the theory of thermal radiation from the atomistic and statistical viewpoint, under the influence of the steady success of atomism.

§4. Stefan-Boltzmann’s law

In 1879 the fourth power law of the total radiation energy was presented by Stefan, which was derived theoretically by Boltzmann in 1884.

3) K. Amano, ibid., pp. 10, 48.
This law states that the density $u$ of the total radiation energy is proportional to the fourth power of the absolute temperature $T$:

$$u = aT^4. \quad (3)$$

Some confusion arose because Stefan supposed that this law would be valid in general without taking into account the condition of black body.\(^4\) Boltzmann derived with the method of thermodynamics the relation between the radiation pressure and the temperature, starting from Bartoli's work mentioned above, which showed that the radiation pressure can be treated like the gas pressure in no contradiction with the second law of thermodynamics. Boltzmann derived also the explanation of the radiation pressure, by applying to it Maxwell's revolutionary electromagnetic theory of light.\(^5\)

\(^4\) K. Amano, ibid., p. 19.
\(^5\) The radiation pressure $p$ is represented by the component of Maxwell's stress normal to a plane, a plane normal to the $x$-axis, say. The component is expressed as

$$p_{xx} = \frac{1}{8\pi} (E_x^2 - E_y^2 - E_z^2) + \frac{1}{8\pi} (H_x^2 - H_y^2 - H_z^2),$$

in terms of the electric and magnetic fields $\mathbf{E}(E_x, E_y, E_z)$ and $\mathbf{H}(H_x, H_y, H_z)$, respectively. Since the pressure is given by the average over the time $\bar{p}_{xx}$ of $p_{xx}$, one makes use of the relations

$$\bar{E}^2 = \bar{H}^2 = \frac{1}{3} \bar{E}^2, \quad \bar{E}^2 = \bar{H}^2 = \frac{1}{3} \bar{H}^2,$$

which hold for the respective field quantities on the average, to get

$$\bar{p}_{xx} = -\frac{1}{3} \cdot \frac{1}{8\pi} (\bar{E}^2 + \bar{H}^2) = -\frac{1}{3} u = -p.$$

The negative sign before $p$ means that the pressure is in the opposite direction to the normal.

The last equation thus gives the relation between the radiation pressure and the density of the total radiation energy. One then puts this relation into the second law of thermodynamics

$$\frac{dU + pdV}{T} = dS,$$

where $U$, $V$ and $S$ denote the total radiation energy, the total volume and the entropy respectively. Because $U = V u$, one has

$$dU = d(Vu) = V \frac{du}{dT} dT + udV.$$

Inserting this and $p = u/3$ into $dS$ given above, and using the condition of complete differential

$$\frac{\partial}{\partial V} \frac{\partial S}{\partial T} = \frac{\partial}{\partial T} \frac{\partial S}{\partial V},$$
In the way the theoretical basis was made clear of the relation between the total radiation energy and the temperature. Kirchhoff showed that the nature of black body radiation is determined by a universal function \( F(\lambda, T) \) only of the wavelength and the temperature. He indicated the importance of finding out this function, saying in the following way:

Though difficulties lie in the experimental determination of it, we may have a well-founded hope of knowing it by means of experiment. This is because it will also be almost doubtless a simple function, like all the functions hitherto known that are independent of the nature of individual bodies have simple forms. Only after the solution of this problem, the whole effectiveness of the proposition proved above will become clear.\(^6\)

On this function Stefan-Boltzmann’s law imposes the condition

\[
\int_0^\infty F(\lambda, T)d\lambda = \text{const.} \cdot T^4. \tag{4}
\]

§5. **Wien’s displacement law**

There arises then the question of finding the dependence of this universal function on wavelength, that is, the distribution law. Before the complete discovery of this law, a relation between wavelength and temperature was found by Wien\(^7\) in 1893 as a condition of this function. This relation was derived from a thermodynamical consideration of the change in wavelength, which is caused to black body radiation by a moving reflective wall on account of the Doppler effect. It is called Wien’s displacement law, and states that, in the normal spectrum of black body radiation, every wavelength changes with temperature holding the product of wavelength by temperature constant. In other words, if \( \lambda \) changes into \( \lambda_0 \) as \( T \) changes into \( T_0 \), one has

\[
T\lambda = T_0\lambda_0 = \text{const.} \tag{5}
\]

one gets

\[
\frac{du}{dT} = \frac{4u}{T},
\]

from the integration of which one obtains

\[
u = aT^4.
\]


To show this, Wien considers, first, in the case of a box with perfectly reflective walls enclosing radiation, to move the walls with velocities small compared with the light velocity so that the density of the radiation energy may remain homogeneous over the whole space within the box. With the use of a clever thought experiment, he proves by means of the second law of thermodynamics that, if the volume of the box which contains initially black body radiation at a temperature is changed, and the density of the total radiation energy in it is made equal to that of black body radiation at another temperature, the spectrum of the density of the radiation energy in the box is the same as that of black body radiation at this temperature over the whole range of the wavelength. That is, the radiation in the box preserves the distribution law of black body radiation.

Wien treats then the case in which one of the walls is a piston moving with a velocity \( v \) small compared with the light velocity \( c \), to calculate the change in wavelength, or the change in period caused by reflection in the direction normal to the wall. Let \( \tau \) and \( \tau' \) be the periods before and after the reflection, respectively.

According to the Doppler effect one has the relation

\[
\tau' = \frac{c - 2v}{c} \tau.
\]  

(6)

Because of \( \tau = \lambda/c \) and \( \tau' = \lambda'/c \), this gives the relation

\[
\lambda' = \frac{c - 2v}{c} \lambda.
\]  

(7)

Now, let the dependence of the density of radiation energy in the box on wavelength be given by the function \( \phi(\lambda) \), so that the density of radiation energy included between \( \lambda \) and \( \lambda + d\lambda \) is expressed as \( \phi(\lambda)d\lambda \). The distribution of radiation energy after \( n \) times of reflection \( f_n(\lambda) \) is calculated to be

\[
f_n(\lambda) = \phi\left(\lambda + \frac{nl}{3}\right),
\]  

(8)

where \( l \) is the shortening of wavelength in a single reflection. If the box is assumed to be a rectangular parallelepiped, and its initial size in the direction of the movement of the piston is denoted by \( a \), the light is known to make \( n \) round runs in the box while the distance \( x \) covered by the piston increases by \( dx \), where
Quantum of Radiation

\[ n = \frac{dx}{2(a - x)} \cdot \frac{c}{v}. \]  
(9)

The use of this relation for the wavelength after \( n \) times of reflection

\[ \lambda_n = \lambda + nl = \left( \frac{c - 2v}{c} \right)^n \lambda, \]  
(10)

results in

\[ nl = -\frac{dx}{a - x} \lambda, \]  
(11)

if infinitesimals of the second order are neglected. In other words, one has

\[ f_n(\lambda) = \phi \left[ \lambda - \frac{dx}{3(a - x)} \lambda \right]. \]  
(12)

This equation is independent of \( n \), and can therefore be written as

\[ f(\lambda) = \phi(\lambda + d\lambda), \]  
(13)

in terms of

\[ d\lambda = -\frac{dx}{3(a - x)} \lambda. \]  
(14)

The integration of this equation gives

\[ \lambda = \sqrt[3]{\frac{a - x}{a}} \lambda_0, \]  
(15)

where \( \lambda_0 \) denotes the wavelength when the piston is at its initial position.

Wien applies next the second law of thermodynamics for adiabatic processes,

\[ dU = -pdV. \]  
(16)

Since the movement of the piston is one-dimensional, the energy density is given by

\[ u = \frac{U}{a - x}, \]  
(17)

where \( U \) is the total energy in the box when the piston has covered the distance \( x \), the area of the piston being taken to be unitary. The energy density \( u \)
increases with the movement of the piston, on account of the decrease in the volume, on one hand, and of the work done by the piston, on the other hand:

\[
\frac{du}{dx} = \frac{\partial u}{\partial x} \frac{dU}{dx} + \frac{\partial u}{\partial x} dx.
\]  

(18)

Because the pressure acted upon the piston is given by

\[ p = \frac{1}{3} u, \]

(19)
as is shown above, one has

\[
\frac{dU}{dx} dx = \frac{1}{3} u dx.
\]

(20)

With the use of this relation, one gets

\[
du = \frac{4}{3} \frac{u}{a-x} dx,
\]

(21)

from which one obtains by integration

\[
u = 3 \sqrt{\left(\frac{a-x}{a}\right)^4} u_0,
\]

(22)

where \( u_0 \) is the initial energy density.

From Eqs. (15) and (22), it results that

\[
\frac{u}{u_0} = \frac{\lambda_0^4}{\lambda^4}.
\]

(23)

Thus, one sees from the description given above, that the distribution law of the density \( u \) of the energy radiated from a body at a temperature \( T \) is identical with that of the energy density of the radiation which is obtained by compressing the radiation of the energy density \( u_0 \) at a lower temperature \( T_0 \). Now, according to Stefan-Boltzmann's law one has

\[
\frac{u}{u_0} = \frac{T^4}{T_0^4}.
\]

(24)

This gives, together with Eq. (23),

\[
T \lambda = T_0 \lambda_0 = \text{const.},
\]

(25)

that is nothing but Wien's displacement law.
Quantum of Radiation

The proof given above follows the one-dimensional treatment in Wien's first paper. As was done by Planck, the proof can be given in a general way by considering reflections of light at a angle to the piston.

Wien's displacement law can also be derived more generally to determine the universal function \( F(\lambda, T) \) in a more particular form. Between the energy density \( u_\lambda \) considered as a function of wavelength \( \lambda \) and the energy density \( u_\nu \) considered as a function of frequency \( \nu \), there is the relation

\[
    u_\lambda d\lambda = -u_\nu d\nu.
\]

Because \( \lambda = c/\nu \), one has

\[
    u_\lambda d\lambda = u_\nu \frac{c}{\lambda^2} d\lambda.
\]

The universal function is known to take the form

\[
    \frac{\lambda}{c} F(\lambda, T) = \frac{c^2}{\lambda^5} \varphi \left( \frac{c}{\lambda T} \right) = \frac{\nu^5}{c^2} \varphi \left( \frac{\nu}{T} \right).
\]

2. Atomistic Investigations of Black Body Radiation

§1. Application of atomism

The studies of thermal radiation described in the previous section were based on Maxwell's theory of electromagnetics and the second law of thermodynamics. In order to make a further step toward the distribution law, it was necessary in addition to give atomistic consideration. This could be developed on the basis of atomistic viewpoint, which was advanced by Boltzmann and others to gain victories step by step over the phenomenological and empirical energetics.

As was described in Wien's paper, in which his famous distribution law was presented, any change in wavelength of black body radiation takes place without consumption and absorption of energy because black body radiation

\[ 8\)From this one has for the wavelength \( \lambda_m \) at which the energy density becomes maximum the relation \( \lambda_m T = \text{const.} \)

\[ 9\)W. Wien, Wied. Ann. 58 (1896), 662. \]
corresponds to a state of maximum entropy, and the energy distribution in the spectrum of black body radiation can well be determined from the condition of maximum entropy, if some process of increasing the entropy is known. In the studies described above, though it is possible to give the entropy of radiation of given intensity and wavelength, "there is no physical process to bring about changes in color in a clear-cut way, so that it is impossible to determine the energy distribution without some assumption". This viewpoint had already been taken by V. A. Michelson in Moscow in an attempt made in 1887, before Wien's work on the displacement law. Michelson (1) considered that the continuity of the spectrum of a solid body was caused by the complete disorder of oscillations of the constituent molecules of the solid body, and applied Maxwell's law of velocity distribution to the molecules, (2) assumed that the period $\tau$ of the oscillation excited by a molecule was related to its velocity $w$ by the relation $\tau = 4a/w$, $a$ being a constant, and (3) supposed that the intensity of the radiation emitted by molecules was proportional to the number of the molecules having the same frequency as that of the radiation, and to an unknown function of the temperature and the kinetic energy of such a molecule which was assumed to be given by a power of $w^2$.

From these assumptions Michelson obtained for the energy density per unit wavelength the formula

$$u_\lambda = B_1 T^{3/2} e^{-C/T \lambda^2} \lambda^{-6},$$

where $B_1$ and $C$ are constants.\textsuperscript{11} This formula gives the relation $\lambda_m^2 T = \text{const.}$, for the wavelength $\lambda_m$ at which the energy density becomes maximum. At that time this relation was considered to be right, because the experimental result was only that by Langley which turned out incorrect. In 1883 H. F. Weber obtained with a purely empirical method the experimental formula

$$u_\lambda = C \frac{1}{\lambda^2} e^{aT - 1/b^2 T^2 \lambda^2},$$

where $C$, $a$ and $b$ are constants. From this formula he got the relation $\lambda_m T = \text{const.}$, showing that Michelson's theoretical formula did not fit the experimental result. This relation is the one that was confirmed theoretically by Wien in 1893 as the displacement law.

\textsuperscript{10}V. A. Michelson, Journ. d. Phys. 6 (1887), 467.

\textsuperscript{11}Cf. also Sec. 1, §1, Chap. 1, Vol. II of the present book.
Though the electromagnetic theory as the theory of emission and absorption of radiation was laid its foundation by the experimental proof of electromagnetic wave done by Hertz in 1888, it could not immediately be applied to the theory of thermal radiation, so that Wien’s distribution law should start by being given rather unfamiliar assumptions as those in Michelson’s theory mentioned above. In other words, the development of the theory of thermal radiation was accomplished by building theories on suitable assumptions from the atomistic viewpoint, before the results derived from the classical theory in strictly logical ways were shown to be incompatible with experiments. It was made clear in later studies done by physicists, from Rayleigh to Einstein and Jeans, that the classical theory was not applicable to the phenomena of thermal radiation. This may be one of the reasons, that the quantum theory came to be accepted as revolutionary by physicists in general for the first time, at such a time as 1911 when the first Solvay conference was held.

§2. **Wien’s distribution law**

As we have described above, Wien derived in 1896, the distribution law of black body radiation from an atomistic viewpoint. In this work, Wien makes use of Maxwell’s law of velocity distribution as Michelson. Wien considers, however, restricting the excitation of radiation by Stefan-Boltzmann’s law and Wien’s displacement law, which are obtained with the use of the electromagnetic theory and thermodynamical method, because about it there is so little knowledge as to make any reliable assumption. Wien says that “the still remaining hypothesis can as usual be allowed to return to uncertainty in the theoretical foundation”, but its results can be extensively compared with experience, and “the confirmation or refusal by experience will decide the correctness or incorrectness of the hypothesis and in this respect it will be useful for further improvement of the molecular theory”. These words of Wien are very noticeable.

Because black body radiation is realized in a cavity enclosed by perfectly reflecting walls, when the radiation in it is emitted by a gas having a finite absorptive power for all wavelengths, Wien applies Maxwell’s law of velocity distribution to the gas molecules, according to which the number of the molecules having a velocity between \( v \) and \( v + dv \) is proportional to \( v^2 e^{-\nu^2/\alpha^2} dv \), \( \alpha \) being a constant related to the mean velocity, or the absolute temperature. Wien says that, though it is entirely unknown how the frequency of the radiation
emitted by a molecule of velocity \( \nu \) is related to the state of the molecule, "the view that the electric charge of a molecule excites electromagnetic waves may be assumed in general at present".

Wien considers first, as an assumption, that every molecule emits radiation of a certain wavelength and intensity, both of which depend only on its velocity. Though a variety of arbitrary suppositions can be made, he chooses to make a simple and general supposition as far as possible. Because the wavelength \( \lambda \) of the radiation emitted by a molecule of velocity \( \nu \) is a function of \( \nu \), \( \nu \) is inversely a function of \( \lambda \). Wien assumes thus that the intensity \( \varphi_\lambda d\lambda \) of the radiation of the wavelengths lying between \( \lambda \) and \( \lambda + d\lambda \) is proportional to

1. the number of the molecules emitting the radiation of these wavelengths,
2. a function of the velocity \( \nu \), or of the wavelength \( \lambda \).

\( \varphi_\lambda \) is therefore given by

\[
\varphi_\lambda = g(\lambda)e^{-f(\lambda)/T},
\]

where \( g(\lambda) \) and \( f(\lambda) \) are unknown functions. We see that \( \nu \) and hence the molecules are eliminated in this equation by the supposition of \( \nu \) being a function of \( \lambda \).

Now, from the comparison of Eq. (30) with the universal function

\[
F(\lambda, T) = \frac{c^2}{\lambda^5} \varphi \left( \frac{c}{\lambda T} \right)
\]
given in Eq. (27), which is derived from Wien’s displacement law, it turns out that

\[
\frac{f(\lambda)}{T} = \frac{C}{\lambda T},
\]

where \( C \) is a constant. On the other hand, because of Stefan-Boltzmann’s law, there holds the relation

\[
\int_0^\infty g(\lambda)e^{-C/\lambda T} d\lambda = \text{const} \cdot T^4,
\]

from which \( g(\lambda) \) can be determined by means of the method of undetermined coefficients. In other words, by inserting the power series expansion of \( g(\lambda) \)
in $\lambda$ into Eq. (32), and comparing the both hand sides with the use of the dimension-less variable $y = C/\lambda T$, one gets

$$T^{-n-1} = T^4,$$

which means that $n = -5$. Thus, one has for $g(\lambda)$

$$g(\lambda) = \frac{\text{const.}}{\lambda^5} \quad (33)$$

and therefore for $\varphi_\lambda$

$$\varphi_\lambda = \text{const.} \frac{e^{-C/\lambda T}}{\lambda^5}. \quad (34)$$

This distribution formula takes the same form as that found experimentally at that time by Paschen. However, Paschen’s experiment was not determinate because it did not satisfy the condition of black body.

§3. **Planck’s effort**

From nearly that time experiments on black body radiation became prosperous in connection with demands by German industry, getting higher accuracy, as is shown by Amano in his book cited above. In particular, in 1895 Wien and O. Lummer showed from Kirchhoff’s law that the cavity radiation realizes the ideal condition of black body radiation, and that the experiments done in earlier times were defective in respect to the condition of black body. After their studies experiments on black body radiation came to have a reliable foundation. Thus, in 1898 O. Lummer and F. Kurlbaum accomplished the black body electric furnace. In the same year Kurlbaum published his measurement of the absolute intensity of black body radiation, and also the study of black body radiation with the use of remnant infrared rays. In 1899 the famous spectroscopic measurements of black body radiation by Lummer and Pringsheim and by Paschen were published.

Planck started his efforts at that time. In order to find the universal function proposed by Kirchhoff, Planck considered instead of making an unfamiliar assumption as Wien did, that black body radiation was realized through the equilibrium, which was arrived at in a cavity by electromagnetic waves emitted and absorbed by linear oscillators, because that time was just after Hertz’s experiment which made clear that a linear oscillator emits and absorbs electromagnetic waves. Planck thought that for this purpose it was sufficient to
set up a theory on the basis of electrodynamics etc. as they were at that time without the need of being restricted particularly. In 1895 Planck started his study of the problem of emission and absorption of electromagnetic waves. In this way, he succeeded in getting a general relation between the energy of an oscillator and the energy of radiation, that was independent of the oscillator. His theory could not take into account the entropy, however, because it was a theory of reversible processes as long as it was a pure electromagnetic theory. When this point was indicated by Boltzmann, Planck introduced the idea of natural radiation, in the same way as Boltzmann introduced the molecular disorder into the second law of thermodynamics. That is, Planck supposed that the energy of radiation was distributed completely in disorder among the individual partial oscillations constituting the radiation. On this supposition, however, he could not achieve his purpose yet. Then, he conversely began to treat the problem from the side of thermodynamics, and became aware of relating the entropy of an oscillator to its energy (about in 1899). But, he made no more progress, because "Planck inclined so much toward phenomenological thinking that he did not made any detailed consideration of the relation between entropy and probability", paying attention only to experimental data at that time.

As Amano shows in his book cited above, in 1891 Planck was contrary together with Ostwald, an advocate of phenomenological energetics, to Boltzmann, an advocate of atomism, but in 1895 Planck conversely refuted with Boltzmann, a materialist, Ostwald's opinion of "overcoming scientific materialism". Planck attacked the energetics as in reality a baseless assertion in spite of its apparent universality, for the reason that "it could not bring forth at all any new result to be verified by experience" though it would give birth to no contradiction.

§4. **Foundation by Planck of Wien's distribution formula**

About in 1899 Planck was collecting experimental data for the distribution law, but the experiments at that time were all to support in a certain degree Wien's distribution formula. Only the experiments on remnant infrared rays showed that Wien's distribution formula would not be valid. This possibility was not

---

13) K. Amano, *ibid.*, p. 53. Amano's description of these circumstances is in particular excellent. As a materialist, he gives really a sharp critique of Machism.
affirmed with certainty, however, because these experiments were difficult ones. Because of the fact, in addition to the prevailing belief in Wien’s distribution formula, that Planck succeeded in deriving Wien’s formula from Wien’s displacement law together with a simple consideration, according to which the inverse of the second derivative of the entropy of an oscillator with respect to its energy was proportional to the energy, Planck thought that Wien’s distribution formula was a correct law. Even in 1900, when it was finally made clear through the experiments done step by step by Lummer-Pringsheim and others that Wien’s formula did not fit the experimental data, Planck persisted in Wien’s formula.

In other words, in a paper in 1900, Planck\(^{14}\) takes the following method. Though the law of energy distribution can be immediately determined, if the entropy \(S\) of a linear resonator oscillating in response to radiation is known as a function of its energy of oscillation, the law of increase of entropy does not suffice alone to determine this function completely. Planck derives it by calculating in two ways the infinitesimal increase in the entropy of a system of \(n\) identical resonators in a stationary radiation field.

Now, suppose that a resonator in a stationary radiation field has the energy

\[
U = U_0 + \Delta U, \tag{35}
\]

which is different from its stationary value \(U_0\) by \(\Delta U\). If its energy changes by \(dU\) during the time \(dt\), the change in its entropy is given by

\[
dS = \frac{dS}{dU} dU. \tag{36}
\]

The insertion of Eq. (35) into Eq. (36) results in

\[
\frac{dS}{dU} = \left( \frac{dS}{dU} \right)_0 + \left( \frac{d^2 S}{dU^2} \right)_0 \Delta U + \cdots. \tag{37}
\]

The entropy \(S_t\) of the whole system of the radiation field and the resonator is given by

\[
S_t = S + \int s dV, \tag{38}
\]

where \(s\) is the density of the entropy of the radiation field, whose integration over the whole space gives the entropy of the radiation field. Consider that \(S_t\)

\(^{14}\)M. Planck, Ann. d. Phys. 1 (1900), 719.
increases by
\[ dS_t = dS + d \int sdV, \]
if the resonator changes its energy by \( dU \) when its energy \( U \) has the infinitesimal increase \( \Delta U \) over its stationary value \( U_0 \). With the use of the equations of the interaction between the resonator and the radiation field, that is, the electromagnetic field, Planck obtains the relation
\[ dS_t = dU \Delta U \frac{3}{5} \frac{d^2 S}{dU^2}. \]
(40)
Thus, \( dS_t \) depends only on \( dU, \Delta U \) and \( U \). Moreover, it is proportional to \( dU \) and \( \Delta U \). Because \( dU \) and \( \Delta U \) are shown to be opposite each other in their sign from the relation between the resonator and the radiation field, Planck puts
\[ \frac{3}{5} \frac{d^2 S}{dU^2} = -f(U), \]
(41)
where \( f(U) \) is a positive function of \( U \). \( dS_t \) is accordingly expressed as
\[ dS_t = -dU \Delta U f(U). \]
(42)
Planck considers the case of \( n \) resonators which make independently of each other the same kind of process. Their total energy \( U_n = nU \) is supposed to differ by an infinitesimal quantity \( \Delta U_n = n\Delta U \) from its stationary value \( U_{0n} = nU_0 \), so that
\[ U_n = U_{0n} + \Delta U. \]
(43)
Their total entropy is the sum of their respective entropies, and is given by \( S_n = nS \). \( U_n \) is then supposed to change by \( dU_n = ndU \) during the time \( dt \).
Now, Planck discusses as follows. “One could for the moment tend towards the supposition that \( S_n \) depends on \( U_n \) in the same way as \( S \) does on \( U \). In that case, it would be possible to get the value of \( S_n \) by putting \( U_n \) instead of \( U \) in the expression of \( S \) taken as a function of \( U \). However, this supposition could not be founded by any consideration of a physical kind, since no physical meaning must actually at all be attached to the quantity of entropy, just as it is the case with the absolute quantity of a potential of force, for example. Rather, only the increase in entropy of the whole system coming in during the time element \( dt \) possesses a definite physical meaning, as this forms the
numerical measure for the irreversibility of the process or for the uncompensatable transformation of work to heat, and some corresponding conclusion can actually be applied to or made on this quantity". He then puts forth his basic idea — "The necessary physical connection between the change in energy and the increase in entropy would hardly be likely to appear understandably, if one does not want to assume that the increase in entropy in the case of the radiation under consideration of the \(n\) resonators is completely determined by their energy \(U_n\), its deviation \(\Delta U_n\) from the stationary value, and the change in energy \(dU_n\) occurring in the time \(dt\), and that as a result of this the quantity of the increase in entropy can then be obtained, if one puts \(U_n\) instead of \(U\) everywhere in the expression" of \(dS_t\) given by Eq. (42).

In this way, Planck extends Eq. (42) to the case of \(n\) resonators. The new \(dS_t\) should be, on the other hand, \(n\) times the old \(dS_t\), because the \(n\) resonators make identical processes at the same time and independently of each other. Planck thus gives the equation

\[
\{ -dU \Delta U f(U) \}_U = -ndU \cdot \Delta U \cdot f(U),
\]

or

\[
dU_n \Delta U_n f(U_n) = ndU \Delta U f(U).
\]

Putting all \(U_n\)’s in this equation to be \(nU\), he gets

\[
f(nU) = \frac{1}{n} f(U).
\]

The solution of this functional equation is

\[
f(U) = \frac{\text{const.}}{U}.
\]

Inserting this into Eq. (41), he has with a constant \(b > 0\)

\[
\frac{d^2S}{dU^2} = -\frac{1}{bU}.
\]

\(b\) depends only on the frequency \(\nu\). The integration of Eq. (48) gives

\[
\frac{dS}{dU} = \text{const.} - \frac{1}{b} \log U,
\]
from which Planck obtains with the use of the second law of thermodynamics $dS/dU = 1/T$

$$U = ae^{-b/T},$$

(50)
a being a coefficient. Wien's displacement law shows that

$$a = Av, \quad b = B\nu,$$

(51)
where $A$ and $B$ are numerical constants. Planck arrives thus at Wien's distribution formula. This theory seemed to Planck to be natural, because in it any special assumption such as Wien's one was included and only the resonators were taken into consideration. Wien's distribution formula seemed to Planck accordingly to be a necessary and natural one.

§5. Rayleigh's critique based on the theory of vibration

The year 1900 was eventful to the theory of thermal radiation. On one hand, many experiments were carried out on thermal radiation, and the experiments on long wavelength rays, especially those on infrared rays, made the deviation of Wien's formula undeniable. On the other hand, Rayleigh\(^{15}\) presented a formula different from Wien's one, by calculating the number of the proper vibrations of the radiation field with frequencies between $\nu$ and $\nu + d\nu$ by analogy with those of air in a cubic box, and by applying the equipartition law of energy to them, and criticized Wien's theory. Rayleigh's formula was derived by Jeans\(^{16}\) later in 1905 strictly from the classical electromagnetic theory.

In the paper cited above, Rayleigh says that "Paschen finds that his observations are well represented" by Wien's distribution formula; "nevertheless, the law seems rather difficult of acceptance, especially the implication that as the temperature is raised, the radiation of given wave-length approaches a limit", indicating just the point to be examined. Rayleigh thus "ventures to suggest a modification of Wien's formula" on the basis of the theory of vibration, which is his specialty. He considers that although for some reason not yet explained the Boltzmann-Maxwell doctrine of the partition of energy, according to which every mode of vibration should be alike favoured, fails in general, it seems possible that it may apply to the graver modes. In the case of

\(^{15}\)Lord Rayleigh, Phil. Mag. 49 (1900), 539.

\(^{16}\)J. H. Jeans, Phil. Mag. 10 (1905), 91.
a string vibrating transversely in one dimensional way, the Boltzmann-Maxwell law states that the energy should be equally divided among all modes, whose frequencies are as 1, 2, 3, ... (in units of the fundamental frequency). Hence, if \( k = 1/\lambda \), representing the frequency, the energy between \( k \) and \( k + dk \) is represented by \( dk \) when \( k \) is large enough.

Passing to the case of three dimensions, and considering for example the vibrations of a cubical mass of air, one has the equation

\[
k^2 = p^2 + q^2 + r^2 ,
\]

where \( p, q, r \) are integers representing the number of subdivisions in the three dimensions. If \( p, q, r \) are regarded as the coordinates of points forming a cubic array, \( k \) is the distance of any point from the origin. Hence, the number of points for which \( k \) lies between \( k \) and \( k + dk \), being proportional to the volume of the corresponding spherical shell, is represented by \( k^2 dk \), and "this expresses the distribution of energy according to the Boltzmann-Maxwell law, so far as regards the wave-length or frequency". Rayleigh concludes thus that "if we apply this result to radiation, we shall have, since the energy in each mode is proportional to \( T \),

\[
Tk^2 dk ,
\]

or, if we prefer it,

\[
T\lambda^{-4} d\lambda" .
\]

Rayleigh makes the suggestion that this may be the proper form than \( \lambda^{-5} d\lambda \), which results from Wien's formula when \( \lambda T \) is great. Now, he says without any reasoning that "if we introduce the exponential factor, the complete expression will be

\[
c_1 T\lambda^{-4} e^{-c_2/\lambda T} d\lambda",
\]

and that "if, as is probably to be preferred, we make \( k \) the independent variable, this becomes

\[
c_1 Tk^2 e^{-c_2k/T} dk" .
\]

His introduction of this exponential factor might be made groundlessly, on account of his idea, described at the beginning of the present subsection, of suggesting a modification of Wien's formula.
§6. Breakdown of the classical theory

Rayleigh's idea consisted in taking the condition of the oscillation of radiation in space into consideration. In the theory of thermal radiation, radiation as an electromagnetic wave was applied for the first time to the calculation of radiation pressure in the derivation of the Stefan-Boltzmann law, and then to the Doppler effect in the derivation of Wien's displacement law. No application of the other laws of radiation field was made, until the calculation of the proper oscillations of radiation in a cavity was made first by Rayleigh, and then in 1905 by Jeans strictly from the viewpoint of the classical electromagnetic theory, that is, a theory of continuous wave. Thus, Rayleigh-Jean's formula is to give the spectral distribution of black body radiation according purely to the classical theory, and expresses the energy density $u_\nu d\nu$ in the frequency interval $d\nu$ as

$$u_\nu d\nu = \frac{8\pi\nu^2}{c^3}kT d\nu.$$  \hspace{1cm} (57)

Rayleigh-Jean's formula contains a great difficulty. In other words, this formula implies that black body radiation has a great deal of energy at high frequencies even if the temperature is very low, so that any body of low temperature should appear brilliant, as is seen from the equation of the total energy$^{17}$

$$\int_0^\infty \frac{8\pi}{\lambda^4}kT d\lambda = \infty.$$  \hspace{1cm} (58)

This difficulty is called the ultraviolet catastrophe. Since Rayleigh made his study from the standpoint of modifying Wien's formula by taking into consideration the vibrations in a cavity, he introduced the exponential factor probably from a molecular theoretical idea, in order to settle the ultraviolet catastrophe and make his formula fit the fact that Wien's formula was in accord with the experimental data for short wavelengths.$^{18}$ It seems that Rayleigh thought that the exponential factor could not be deduced from the classical theory. In this way, about in 1900 it was not thought clearly that the classical electromagnetic theory was in contradiction with black body radiation, even after

$^{17}$This point was pointed out in particular by Einstein, as we shall describe later.

$^{18}$Rayleigh's formula became shown not to fit the reality by Lummer-Pringsheim's experiment.
the proposal of Planck’s theory of energy quantum. This contradiction became decisively recognized in the studies by Einstein and by Jeans in particular.¹⁹)

§7. Planck’s new distribution formula

Because it was made clear that Wien’s formula was off the experimental data for long wavelengths by Lummer-Pringsheim’s work, and in particular by Rubens-Kurlbaum’s result²⁰) of the measurement of remnant infrared rays from rock salt (51 µ in wavelength), Planck²¹) proceeded to reexamine his former thermodynamical consideration. In the new study under consideration, Planck admits that Wien’s distribution formula does not possess a general meaning but has at most the character of a limiting law, and accepts now the opinion that “its extremely simple form owes its origin just to a restriction to short wavelengths or low temperatures”, contrary to the thought in his former study that Wien’s formula should be right because of the simplicity of the thermodynamical equations used for it. About Eq. (45), that is,

\[ dU_n \Delta U_n f(U_n) = n dU \Delta U f(U), \]

which was taken as the starting point in the former study, Planck says “I would consider it well possible, even so not easily understandable and anyway hardly provable, that the expression does not possess generally the meaning ascribed formerly by me to it”.

Then, about the mistake²²) of having simply applied in a formal way the expression \( dU \Delta U f(U) \) derived for a single resonator to the case of \( n \) resonators in Eq. (45), considering that “the values of \( U_n, dU_n \) and \( U_n \) are not sufficient to determine the change in entropy in question, but at the same time \( U \) must also be known”, he states “Following this thought I have finally come to construct quite arbitrary expressions of the entropy, which, though more complicated than Wien’s expression, nevertheless seem to satisfy all the demands of the thermodynamical and electromagnetic theories just as much as that”. Planck starts therefore this time by putting semiempirical equations, instead of basing on theoretical necessities as in the former study.

Among such expressions, one that is next to Wien’s in simplicity makes a special impression on Planck. It is obtained by putting

\[
\frac{d^2 S}{dU^2} = -\frac{\alpha}{U(\beta + U)}.
\]  

(59)

This is the simplest of all the equations giving \(S\) as a logarithmic function of \(U\), which lies near to the probability theoretical calculation. For small values of \(U\) the right hand side of Eq. (59) behaves as

\[
\frac{\alpha}{U(\beta + U)} \rightarrow \frac{\alpha/\beta}{U},
\]  

(60)
tending to the expression to give Wien’s formula. The integration of Eq. (59) together with the second law of thermodynamics gives

\[
\frac{dS}{dU} = \frac{1}{T} = -\frac{\alpha}{\beta} \log \left( \frac{U}{U + \beta} \right) + \gamma,
\]  

(61)

where \(\gamma = 0\) because \(U \rightarrow \infty\) as \(T \rightarrow \infty\). The use of Wien’s displacement law

\[
S = f \left( \frac{U}{\nu} \right),
\]  

(62)

then gives the distribution formula

\[
u_\lambda = \frac{C\lambda^{-5}}{e^{c/\lambda T} - 1},
\]  

(63)

here \(C\) and \(c\) being some constants. Planck stressed that this formula reproduced the observational data as much satisfactorily as the best spectral equations hitherto set up, namely those of Thiesen, of Lummer-Jahnke and of Lummer-Pringsheim.\(^{23}\)

\(^{23}\)Equation (59) becomes easy to be interpreted if its inverse is taken (as in Planck’s Nobel prize lecture). One sees that

\[
1 \left/ \frac{d^2 S}{dU^2} \right. = -bU
\]

leads to Wien’s formula and

\[
1 \left/ \frac{d^2 S}{dU^2} \right. = -\frac{U^2}{a}
\]

to Rayleigh-Jeans’s formula, but

\[
1 \left/ \frac{d^2 S}{dU^2} \right. = -bU - \frac{U^2}{a}
\]
§8. Planck’s effort to found the new formula

The Planck formula of thermal radiation appeared at first in this way as an interpolation formula, quite independently of the idea of energy quantum. That is, it was not determined necessarily and logically with unambiguity, but phenomenologically and descriptively. It should be noted, however, that his formula is never of the restricted sense of a simple Machist description of experience, but is rather based thoroughly on atomism, and rests on the idea of resonators.

It should also be noted that, if Planck had been based on Machist empiricism, any further interpretation of the new formula would have been considered a needless metaphysical one, because a formula describing experience very well had been thus obtained. Then the discovery of the action quantum which dominates the 20th century would not have been gained. Immediately after having set up the new formula, Planck started his effort to find a realistic and physical meaning in it. As is described by Amano,24) “if Planck had been satisfied simply with an empirical formula, the discovery of the action quantum would not have become his scientific achievement”, and “in giving the phenomenon a right more than a simple description” he found a clue to arrive at the great discovery. Here, we see a vivid example of the immediate effectiveness of philosophical methodology on scientific studies. In other words, this shows that Machist idealism is not to open any way to the future, and materialistic efforts only are to cut through a new world. It is hoped that this point will be reconsidered well by those who in recent times take refuge in the viewpoint, that studies in physics are irrespective of their being idealistic or materialistic, though the rightness of materialism could but be admitted.

Thus Planck25) set about founding the distribution formula discovered by him, and went toward Boltzmann’s way of thinking, that is, the consideration of the relation between entropy and probability. This was not only because of its apparent good agreement with the previous results of measurement, but also because of its “very simple expression, which seemed certainly more likely to promise a general interpretation, than every other formula previously brought to suggestion”, as Planck says in his paper cited just above. The idea of the energy element introduced here in a revolutionary way by Planck was the resulting from the summation of the right hand sides of these two equations leads to Planck’s formula.

one that occurred to Boltzmann many times.\textsuperscript{26}) Planck derived suggestions from Boltzmann's study in 1877 of the probability theoretical foundation of entropy, in which gas molecules were treated as taking, in the process of their collisions, discontinuous energies of integer multiples of the energy element $\varepsilon$, for the purpose of making calculation. Planck applied Boltzmann's idea to the resonators which emit and absorb radiation. Though in Boltzmann's study $\varepsilon$ was finally let to tend to zero, Planck obtained his new distribution formula by keeping $\varepsilon$ finite. As is known now, one gets Rayleigh-Jean's formula from Planck's by letting $\varepsilon$ to tend to zero.

\section*{§9. Discovery of energy quantum}

Planck starts, in his paper under consideration, from a statistical treatment of resonators each of which makes a stationary oscillation with a constant energy. In other words, he considers the average value at an instant of the energies of very many resonators of an identical kind, which are in a field of radiation so separately from each other as not to interact directly. The entropy of every resonator is then determined by the ways of distributing a given energy among the resonators at one time, so that it becomes possible to calculate this by applying to the electromagnetic theory of radiation the probability theoretical consideration given by Boltzmann in 1877, which made clear for the first time the meaning of the second law of thermodynamics.

In the paper under consideration, Planck gives the distribution of a given energy among the respective colors in the normal spectrum with the use of a single natural constant, elementarily without relying upon any other theory, and then the temperature of the radiation corresponding to this energy numerically by making use of another natural constant. One should wait for his later book\textsuperscript{27}) about the task "to carry out systematically to the last detail every deduction, which is based on the laws of electromagnetic radiation, thermodynamics and probability calculation".

Let us suppose that there are many monochromatic linear resonators separately from each other at adequate distances, in a diathermic medium enclosed with a reflecting wall. The light velocity in the medium will be denoted as $c$. Of all the resonators, $N$ resonators are supposed to have the frequency $\nu$, $N'$ the frequency $\nu'$, $N''$ the frequency $\nu''$, and so on, where all the $N$'s are assumed to be large. The whole system is supposed to have the total energy $E_t$, \textsuperscript{26})K. Amano, \textit{ibid.}, p. 75.  
\textsuperscript{27})M. Planck, "Theorie der Wärmestrahlung" (Leipzig, 1906).
which is shared by the radiation going on in the medium with the resonators being in oscillation.

Now, the problem is (1) how this energy is distributed among the oscillations of the resonators and among the respective colors of the radiation, and (2) what temperature the whole system has in this case.

To start with, let us consider only the oscillations of the resonators, and give them certain energies, namely the $N$ resonators $\nu$ the energy $E$, the $N'$ resonators $\nu'$ the energy $E'$, the $N''$ resonators the energy $E''$, and so on. Their sum $E_0$ should satisfy

$$E + E' + E'' + \cdots = E_0 < E_t.$$  

$E_t - E_0$ is the energy to be assigned to the radiation. The distribution of the energy is still to be carried out to the individual resonators of each kind. Let us distribute first the energy $E$ among the $N$ resonators of the frequency $\nu$. If $E$ is considered to be an unlimitedly divisible quantity, the distribution is possible in infinitely many ways. "We consider $E$ however — and this is the most essential point of the whole calculation — as composed of a quite definite number of finite identical parts and use for it the natural constant $h$. This constant multiplied by the common frequency $\nu$ of the resonators results in the energy element $\epsilon$, and by the division of $E$ by $\epsilon$ we obtain the number $P$ of the energy elements, which are distributed among the $N$ resonators. If the quotient so calculated is not an integer, one takes for $P$ an integer near by lying". Thus, the relation

$$\epsilon = h\nu$$ (64)

is set up, which Weyl\textsuperscript{28} called "the magical formula".

The problem is then to distribute the $P$ energy elements among the $N$ resonators. The way of making this kind of distribution is the one that Boltzmann called "the complexion". The number of the possible complexions is from the theory of combination

$$\frac{N \cdot (N + 1) \cdot (N + 2) \cdots (N + P - 1)}{1 \cdot 2 \cdots P} = \frac{(N + P - 1)!}{(N - 1)!P!},$$ (65a)

which is approximated as

$$\simeq \frac{(N + P)^{N + P}}{N^N P^P}$$ (65b)

\textsuperscript{28}H. Weyl, "Gruppentheorie und Quantenmechanik" (S. Hirzel, Leipzig, 1928).
with the use of Stirling’s formula \((N! \approx N^N)\). Similar numbers are calculated also for the resonators of the other kinds. The product of all the numbers so calculated gives the total number \(N\) of the possible complexions of distributing the energy among all the resonators.

About the distributions of \(E_0\) to \(E, E', E'' \ldots\) different from that considered above, their respective numbers \(N\) of possible complexions are correspondingly determined. When \(E_0\) is kept constant, such a distribution is uniquely determined that has the number \(N\) of possible complexions greater than any other distribution. It is this distribution that the resonators take in a stationary field of radiation when they have the energy \(E_0\) as a whole. All the quantities \(E, E', E'' \ldots\) are then expressed by the single quantity \(E_0\), and

\[
E/N = U_\nu, \quad E'/N = U'_\nu, \quad E''/N'' = U''_\nu, \ldots .
\]  

(66)
give respectively the stationary values of the energy per resonator of each kind.

As is know from the electromagnetic theory of radiation, the space density of the energy of radiation in a diathermic medium contained in the frequency interval between \(\nu\) and \(\nu + d\nu\) is given by

\[
u_\nu d\nu = \frac{8\pi \nu^2}{c^3} U_\nu d\nu,
\]

(67)
so that the values of \(U_\nu\) given above determine the energy of radiation contained in the medium. Though \(E_0\) is left arbitrary up to now, it is easily calculated from \(E_t\) because the relation between them is stationary.

Being known the stationary distribution of energy in terms of a constant \(h\), the entropy of the system of the resonators is obtained in terms of another constant — the Boltzmann constant \(k\), by means of the relation

\[
S = k \log N.
\]

(68)

\(S\) is the sum of the entropies of the respective resonators. The second law of thermodynamics

\[
\frac{1}{T} = k \frac{d \log N}{dE}
\]

(69)
gives the temperature corresponding to that distribution of energy, from which Planck’s formula is derived.
Planck\textsuperscript{29}) gives a further explanation. The probability $W$ of the $N$ resonators of having the energy of oscillation $U_N$ as a whole, is proportional to the number $N$ of all the possible complexions of distributing $U_N$ among the $N$ resonators. In other words, a definite complexion is equally probable as any other definite complexion. Thus, Planck puts as Boltzmann did

$$S_N = \log W + \text{const.},$$

or

$$S_N = \log N \quad = k \{ (N + P) \log(N + P) - N \log N - P \log P \}. \quad (70)$$

The use of the relations\textsuperscript{30}) $U_N = NU$ and $U_N = P\varepsilon$ results in

$$S_N = kN \left\{ \left( 1 + \frac{U}{\varepsilon} \right) \log \left( 1 + \frac{U}{\varepsilon} \right) - \frac{U}{\varepsilon} \log \frac{U}{\varepsilon} \right\}. \quad (71)$$

The entropy per resonator $S$ becomes

$$S = k \left\{ \left( 1 + \frac{U}{\varepsilon} \right) \log \left( 1 + \frac{U}{\varepsilon} \right) - \frac{U}{\varepsilon} \log \frac{U}{\varepsilon} \right\}, \quad (72)$$

because of the relation $S_N = NS$.

Stefan-Boltzmann's law and Wien's displacement law give

$$u = \frac{\nu^3}{c^3} f\left( \frac{T}{\nu} \right). \quad (73)$$

On the other hand, there is the formula\textsuperscript{31})

$$u = \frac{8\pi\nu^2}{c^3} U \quad (74)$$

which relates the radiation field to the resonators, as is seen from Eq. (67). Eqs. (73) and (74) give

$$U = \nu f\left( \frac{T}{\nu} \right), \quad (75)$$

\textsuperscript{29}) M. Planck, Ann. d. Phys. 4 (1901), 553.
\textsuperscript{30}) $U_\nu$ and $u_\nu$ are abbreviated for the moment as $U$ and $u$, respectively, since a definite frequency $\nu$ is under consideration.
\textsuperscript{31}) M. Planck, Ann. d. Phys. 1 (1900), 99.
or

\[ T = \nu g \left( \frac{U}{\nu} \right). \]  

(76)

With the use of this expression of \( T \) the integration of the second law of thermodynamics \( dS = dU/T \) results in

\[ S = G \left( \frac{U}{\nu} \right). \]  

(77)

The entropy of any resonator in the field of black body radiation is, therefore, depends on the single variable \( U/\nu \). This is the simplest expression of Wien’s displacement law.

The application of Wien’s displacement law to the entropy \( S \) given in Eq. (72) shows necessarily that the energy element \( \varepsilon \) is proportional to the frequency \( \nu \), so that

\[ \varepsilon = h\nu. \]  

(78)

Eq. (72) gives then

\[ S = k \left\{ \left( 1 + U/\nu \right) \log \left( 1 + \frac{U}{h\nu} \right) - \frac{U}{h\nu} \log \frac{U}{h\nu} \right\}. \]  

(79)

The use of this in the second law of thermodynamics \( dS = dU/T \) leads to

\[ \frac{1}{T} = \frac{k}{h\nu} \log \left( 1 + \frac{h\nu}{U} \right) \]  

(80)

and therefore

\[ U = \frac{h\nu}{e^{h\nu/kT} - 1}. \]  

(81)

Eq. (74) gives thus the law of energy distribution\(^{32}\)

\[ u_{\nu} = \frac{8\pi h\nu^3}{e^3} \frac{1}{e^{h\nu/kT} - 1}, \]  

(82)

or on account of \( d\lambda = cd\nu/\nu^2 \)

\[ u_{\lambda} = \frac{8\pi c\lambda}{\lambda^5} \frac{1}{e^{c\lambda/kT} - 1}. \]  

(83)

\(^{32}\)The suffix \( \nu \) is added again.
In this way Planck notices a natural constant in the energy element and emphasizes that the energy element is a certain objective existence, opening the door to the new world of quantum. He determines the values of the natural constants $h$ and $k$ in the following way.

He integrates first Eq. (82) for $T = 1$ by means of the power expansion in $e^{-h\nu/T}$ to know the total energy emitted per second by the black body of $1 \text{ cm}^3$ at that absolute temperature, getting

$$
\int_0^\infty u_\nu d\nu = \frac{48\pi k^4}{c^3h^3} \times 1.0823. \quad (84)
$$

He deduces the value of this energy with the use of Stefan-Boltzmann’s law, from the energy emitted by a black body at $100^\circ\text{C}$ per second and per cm$^3$ into the air at $0^\circ\text{C}$ measured by F. Kurlbaum in 1898,\(^{33}\) to obtain

$$
\frac{k^4}{h^3} = 1.1682 \times 10^{15}. \quad (85)
$$

Planck inserts the value of the product $\lambda_m T$ determined by Lummer-Pringsheim’s experiment into the relation

$$
\lambda_m T = \frac{ch}{4.9651k}, \quad (86)
$$

which is derived about the wavelength $\lambda_m$ at which $u_\lambda$ for the temperature $T$ takes its maximum value, and obtains

$$
\frac{h}{k} = 4.866 \times 10^{-11}. \quad (87)
$$

From Eqs. (85) and (86) he gets

$$
h = 6.55 \times 10^{-27} \text{ erg} \cdot \text{sec},
$$

$$
k = 1.346 \times 10^{-16} \text{ erg}/\text{deg}.
$$

Planck calculates also the number of molecules $R/k$ per gram-molecule, where $R$ is the gas constant (being $8.31 \times 10^7$ for $O = 16$), with the use of the value of $k$ given above, obtaining

$$
\frac{R}{k} = 6.175 \times 10^{23}.
$$

\(^{33}\)Translator’s note: The description of Kurlbaum’s work given here in the present translation is slightly more detailed than that given in Taketani’s original book.
3. Einstein’s Light Quantum

§1. Photoelectric effect

The great discovery by Planck described in the preceding section was not taken up for about five years afterwards. Planck himself could not develop his very idea on other fields of study. This was because Planck’s quantum hypothesis was proposed most of all about resonators, saying that any resonator could not take energies other than integer multiples of its energy element $\varepsilon = h\nu$. In order to develop this idea on other phenomena, the knowledge of the substantialistic structure of the matter was not yet enough. Planck’s resonators were no more than abstract Hertz’s linear oscillators. At that time Thomson’s atomic model was not even presented. In this connection, it may be noted that Lorentz introduced a magnetic field to Hertz’s oscillators and gave a brilliant explanation of the Zeeman effect. However, the application of the quantum hypothesis to this kind of problem could not but be put off until the Bohr theory in 1913. It can be said that this was caused chiefly by the lack of substantialistic knowledge.

Paying attention mainly to resonators, Planck applied the idea of the energy quantum to them, but left the radiation considered only to be in equilibrium with the quantum resonators. About this problem the second great step was advanced by Einstein in 1905. Einstein applied the idea of quantum to the field of radiation, and developed, in opposition to the undulatory theory of light, the corpuscular theory of light by substantialistic treatment of light. For this end, the phenomenon of black body radiation, which is a kind of statistical phenomenon as in Plank’s treatment, did not suffice. In other words, such an elementary process was needed as to be able to bring more substantialistic knowledge. It was just the photoelectric effect.

Hertz discovered in 1887 that, when ultraviolet rays emitted by a electric spark hit the cathode of the spark gap, the electric discharge became easier to occur. Hallwacks discovered in the next year 1888 that, when light of short wavelength struck on the surface of a metallic body, the metallic body became positively charged. Then, in 1889, Elster and Geitel observed the same kind of phenomenon, by irradiating the amalgam of alkali metal (K, Na), Zn, Al, etc. with visible light and ultraviolet rays. This was the discovery of the photoelectric effect.

The study of the photoelectric effect advanced rapidly together with that of X-rays, because Röntgen found in 1897 that such rays as had the same kind of
effect were emitted from matter irradiated with X-rays, after his discovery of X-rays in 1895. These secondary rays were studied by Benoist, Perrin (1897), Langevin, Townsend, Sagnac et al., (1899), and it was found that their main parts had a character entirely different from that of primary X-rays, namely, that they were completely absorbed by air of several mm thickness. Dorn confirmed in 1900 that these secondary rays were negatively charged particles, from the fact that they were deflected by a magnetic field in the same direction as cathode rays. Curie and Sagnac arrived at the same conclusion in 1900 by measuring directly their electric charges.\(^{34}\) Already in the first stage of the study of these secondary rays, a remarkable similarity was found between the emission of electrons by X-rays and that by visible light and ultraviolet rays. This was noticed for the first time by Sagnac. The similarity consisted in the fact that the greatest velocity of the electrons of any secondary ray was independent of the intensity of the primary ray, dependent only on the hardness or the wavelength of the primary ray, and almost no electron was emitted with red and infrared rays. Quantitative investigations on this subject were made by Lenard\(^{35}\) in 1902. It was pointed out that this meant a considerable difficulty for the theory of radiation based on the classical electromagnetic theory. This formed the basis of Einstein's theory.

\section*{§2. Einstein's critique of the classical theory}

Einstein\(^{36}\) paid attention thus to radiation and developed the quantum theory from a viewpoint different from Planck's. In the paper now cited, Einstein considers first that there is a fundamental difference between the way of thinking of gas and ponderable particles, and that of the Maxwell theory of electromagnetic process. In other words, while the state of a body is determined completely by the positions and velocities of a very great but finite number of atoms and electrons, continuous space functions are used for the determination of the electromagnetic state in space, so that a finite number of quantities are not sufficient for its full fixing. According to the Maxwell theory the energy is, in the case of pure electromagnetic phenomena such as light, to be grasped as a continuous space function, while the energy of a ponderable body is expressed

\(^{34}\)It was in 1907 by Bestelmeyer, however, that their electric charges were shown from their deflections in magnetic and electric fields, to have almost the same values as those obtained by Kaufmann for cathode rays.

\(^{35}\)P. Lenard, Ann. d. Phys. 8 (1902), 169.

as the sum of the energies of the atoms and electrons. The energy of a ponderable body cannot be divided into arbitrarily many, arbitrarily small parts, while the energy of a ray of light sent out from a point source is, according to the Maxwell theory (or in general to every undulatory theory), distributed continuously over more and more increasing volume.

The undulatory theory of light with the use of continuous pace functions has proved itself satisfactory for the representation of pure optical phenomena. Optical observations refer however to average values in time, but not to momentary values. In spite of the complete confirmation of the undulatory theory about the phenomena of diffraction, reflection, refraction, dispersion etc., the theory of light with the use of continuous space functions leads to contradictions with the experience about the phenomena of emission and absorption of light.

Einstein sees so clearly for the first time these contradictions, and comes to the idea that, about the black body radiation, photoluminescence, generation of cathode rays by ultraviolet light and other kinds of emission and absorption of light, the energy of light is better understood as being distributed discontinuously in space. In other words, he stresses that “in the spreading of a ray of light going out from a point, the energy is not distributed over larger and larger growing space, but composed of a finite number of energy quanta localizing in space points, which move without dividing into parts and can be absorbed and generated only as wholes”.

§3. Derivation of the radiation formula from the classical theory

Einstein examines thus the problem of black body radiation first. That is, he starts from the Maxwell theory and the theory of electron, and shows that they are in contradiction with black body radiation. This contradiction was not made clear enough in Planck’s study. Einstein supposes that in a space enclosed with completely reflecting walls there are a number of gas molecules and electrons, which are freely movable and exert on each other conservative forces in their mutual approach, i.e. can collide with each other as gas molecules in the kinetic theory of gases. This is equivalent to the assumption that the average kinetic energies of gas molecules and electrons are equal to each other in temperature equilibrium. It is the same as that used by Drude in his theory of thermal and electric conductivities. A number of electrons are in
addition supposed to be bound to their respective points far remote from each other in space, by forces directing to the respective points and proportional to the respective distances from them. These electrons should enter into conservative interaction with the free molecules and electrons. Einstein calls the electrons bound to the space points "resonators". They send out and absorb electromagnetic waves of a definite period.

The radiation in the space under consideration, which is found in the case of dynamic equilibrium on the basis of the Maxwell theory, is identical with black body radiation, at least if resonators of all frequencies considered are regarded as existing.

Einstein ignores for the time being the radiation emitted and absorbed by the resonators, and asks the condition for the dynamic equilibrium corresponding to the collisions of molecules and electrons. The kinetic theory of gases gives the condition for their average energies. Decomposing the motion of each resonator and electron into three oscillating motions perpendicular to each other, one finds for the average value $\bar{E}$ of such a linear oscillating motion, as is mentioned before,

$$\bar{E} = \frac{R}{N} T,$$

(88)

where $R$ is the absolute gas constant and $N$ the number of molecules per gram equivalent.$^{37}$ $\bar{E}$ is $2/3$ times the average kinetic energy of a free monoatomic molecule, because the time average values of the kinetic and potential energies are equal to each other. If it would happen, for some reason — in the present case by the processes of radiation —, that the energy of a resonator has an average value larger or smaller than $\bar{E}$, the collisions with the free electrons and molecules would lead to an interchange of energy different on average from zero with the gas. Dynamic equilibrium is therefore possible only when every resonator has the average energy $\bar{E}$.

Einstein makes a similar consideration regarding the interaction of the resonators with the radiation as existing in the space. Planck$^{38}$) derived the condition of dynamic equilibrium for the present case, on the assumption that the radiation can be considered a thinkably most disordered process. It is Eq. (74) given in the last section

$^{37}$Though $R/N = k$ is the Boltzmann constant, Einstein does not use particularly $k$.

\[ \bar{E}_\nu = \frac{c^3}{8\pi\nu^2} u_\nu, \]  

(89)

where \( \bar{E}_\nu \) stands for \( U_\nu \) used there.

The energy of the radiation of frequency \( \nu \) should neither diminish nor increase as a whole, so that there must hold the relations

\[ \frac{R}{N} T = \bar{E} = \bar{E}_\nu = \frac{c^3}{8\pi\nu^2} u_\nu, \]

(90)

\[ u_\nu = \frac{R 8\pi\nu^2}{N} \frac{T}{c^3}. \]

(91)

Einstein arrives thus at Rayleigh-Jean's formula.\(^{39}\)

Einstein concludes that "this relation found as the condition of dynamic equilibrium not only lacks the agreement with experience but also means that nothing can be said of a particular distribution of energy between ether and matter". He shows also that the wider the range of frequency is chosen the greater the energy of radiation becomes, and one has in the limit

\[ \int_0^\infty u_\nu d\nu = \frac{R 8\pi}{N} \frac{T}{c^3} \int_0^\infty \nu^2 d\nu = \infty, \]

(92)

as we have mentioned in the last section.

§4. Einstein's consideration of Planck's theory

Einstein discusses then the determination by Planck of the elementary quantum, and says that it is to a certain degree independent of the theory of black radiation put forward by Planck. In this connection, it is interesting to note that Einstein says about this point in his paper\(^{40}\) published in the next year as follows — "At that time it seemed to me as if Planck's theory of radiation formed an opposite in certain relation to my work. New considerations, which are communicated in the present work, showed me however, that theoretical basis, on which Planck's theory of radiation rests, differs from the basis, which would arise from the Maxwell theory and the theory of electron, to be precise, that the Planck theory makes use implicitly of the hypothesis of light quantum just mentioned".

\(^{39}\)Einstein makes no reference to Rayleigh's derivation of this formula.

§5. **Two limits of Planck’s formula**

Now, in the paper\(^3\) in 1905, Einstein examines Planck’s formula from a somewhat different side. That is, he compares its limits for long and short wavelengths with Rayleigh-Jean’s and Wien’s formulae respectively. Planck’s formula is

\[
 u_\nu = \frac{\alpha \nu^3}{e^{\beta \nu/T} - 1}, \tag{93}
\]

where

\[
\alpha = 6.10 \times 10^{-56}.
\]

\[
\beta = 4.866 \times 10^{-11}.
\]

For large values of \(T/\nu\), or in the limit of long wave lengths, this formula tends to

\[
 u_\nu = \frac{\alpha}{\beta} \nu^2 T. \tag{94}
\]

In other words, Planck’s formula becomes identical with Eq. (91) derived above from the Maxwell theory and the theory of electron. The comparison of the coefficients of Eqs. (91) and (94) gives

\[
 \frac{R}{N} \frac{8\pi}{c^3} = \frac{\alpha}{\beta}, \tag{95}
\]

that is

\[
 N = \frac{\beta}{\alpha} \frac{8\pi R}{c^3} = 6.17 \times 10^{23}. \tag{96}
\]

Einstein arrives thus, as a restriction on the classical theory, at the conclusion that “the greater the energy density and wavelength of radiation are, the more useful the theoretical bases used by us prove; but for short wavelengths and small radiation densities they break down completely”.

§6. **Meaning of Wien’s formula**

Einstein examines then the bounds of validity of Wien’s formula, and makes clear that within the bounds the radiation is considered equivalent to an aggregate of particles, to which the kinetic theory of gases can be applied.
Taking the case of the radiation which occupies the volume $V$, Einsteins make the following assumption that “the perceptible properties of the existing radiation are completely determined, if the radiation density $\rho(\nu)$ is given for all frequencies”. Without output of work and without supply of heat, radiations of various frequencies are regarded as separable from each other. For the purpose of comparing the radiation with a gas of molecules, the radiation is considered therefore to consist of monochromatic lights. By seeking the condition to be satisfied by the entropy of each of them, it becomes possible to examine Wien’s formula.

From the supposition of the mutual independence of radiations of various frequencies, the entropy of the radiation is represented in the form

$$S = V \int_{0}^{\infty} \varphi(u, \nu) d\nu,$$

where $\varphi$ is a function of $u$ and $\nu$. From the statement made before that by any adiabatic compression of the radiation with the reflecting walls its entropy is not changed, $\varphi$ can be reduced to a function only of a single variable. Einstein investigates, however, how the function $\varphi$ can be determined from the law of black body radiation.

The condition of black body radiation is that the entropy is maximum for a given energy. $u$ is determined to be as such a function of $\nu$ that satisfies this condition. From both the conditions of maximum entropy and of constant energy, it turns out with the use of an undetermined coefficient that $\partial \varphi / \partial u$ is independent of $\nu$. By applying the second law of thermodynamics to the case of the change in temperature of the radiation by a supply of heat, the law of black body radiation for $\varphi$ is obtained to be

\[41\]

\[ \delta \int_{0}^{\infty} \varphi(u, \nu) d\nu = 0, \]

and from the condition of constant energy

\[ \delta \int_{0}^{\infty} ud\nu = 0. \]

From these equations it results that

\[ \int_{0}^{\infty} \left( \frac{\partial \varphi}{\partial u} - \lambda \right) \delta ud\nu = 0, \]
Thus, one can determine from the function \( \varphi \) the law of black body radiation, and reversely from the latter the function \( \varphi \) by means of integration under the condition that \( \varphi = 0 \) for \( \rho = 0 \).

With this preparation Einstein enters now into the examination of Wien’s formula

\[
u = \alpha \nu^3 e^{-\beta \nu/T}.
\]

(99)

As is mentioned already, this formula fits well the experiment for large values of \( \nu/T \), though it does not for long wavelengths. Einstein uses it with this restriction. From it \( 1/T \) is calculated to be

\[
\frac{1}{T} = -\frac{1}{\beta \nu} \log \frac{u}{\alpha \nu^3}.
\]

(100)

From Eq. (98) this is equal to \( \partial \varphi/\partial u \), which is integrated with respect to \( u \) to give

\[
\varphi(u, \nu) = -\frac{u}{\beta \nu} \left\{ \log \frac{u}{\alpha \nu^3} - 1 \right\}.
\]

(101)

Let there be a radiation of the energy \( E \) whose frequency lies between \( \nu \) and \( \nu + d\nu \). The radiation occupies the volume \( V \), so that \( u = E/V \). The entropy

where \( \lambda \) is an undetermined coefficient independent of \( \nu \). \( \partial \varphi/\partial u \) is therefore independent of \( \nu \) for black body radiation, and one has (for \( V = 1 \))

\[
dS = \int_{\nu=0}^{\nu=\infty} \frac{\partial \varphi}{\partial u} dud\nu.
\]

Because \( \partial \varphi/\partial u \) is independent of \( \nu \), this reduces to

\[
dS = \frac{\partial \varphi}{\partial u} \int_{\nu=0}^{\nu=\infty} dud\nu = \frac{\partial \varphi}{\partial u} dE,
\]

where \( dE \) is equal to the supplied amount of heat. The second law of thermodynamics indicates also

\[
dS = \frac{1}{T} dE.
\]

The comparison of these two equations gives

\[
\frac{\partial \varphi}{\partial u} = \frac{1}{T}.
\]

This is the law of determining black body radiation.
of this radiation is
\[ Sd\nu = V\varphi(u, \nu)d\nu = -\frac{E}{\beta\nu} \left\{ \log \left( \frac{E}{V\alpha\nu^3} \right) - 1 \right\} d\nu. \] (102)

For the sake of the subsequent comparison with the kinetic theory of gases, the entropy of the radiation is taken into account only in relation to the volume of the radiation, and the entropy is denoted by \( S_0 \) when it has the volume \( V_0 \). One obtains then
\[ S - S_0 = \frac{E}{\beta\nu} \log \left( \frac{V}{V_0} \right). \] (103)

This equation shows that the entropy of a monochromatic radiation of low enough density varies with the volume according to the same law as the entropy of an ideal gas or a dilute solution. Einstein treats this equation thus found by means of the relation between probability and entropy introduced by Boltzmann.

\section*{§7. Treatment of radiation by means of the kinetic theory of gases}

If it has a meaning to mention the probability of a state, and if further every increase in entropy can be grasped as a transition to a more probable state, the entropy \( S \) of a system is a function of the probability \( W \) of its momentary state. Because, for two independent systems, the probability of the whole system is the product of the probabilities of the respective systems, while the entropy of the whole system is the sum of the entropies of both the systems, the following Boltzmann’s relation is obtained.
\[ S = \frac{R}{N} \log W + \text{const.} \] (104)

If \( S_0 \) is the entropy of a system in its initial state, and \( W \) is the relative probability of the state of the entropy \( S \), one has in general
\[ S - S_0 = \frac{R}{N} \log W. \] (105)

Einstein treats then the following special case. Let there exist in volume \( V_0 \) a number \( (n) \) of moving points (molecules for example). In addition to them there can also exist arbitrarily many other moving points of any kind. About the law of motion of the points under consideration, nothing is assumed other
than that neither any part nor any direction in the space is particular. The number \( n \) of the moving points is further assumed to be so small that their mutual interactions can be neglected.

To the system under consideration, which can be an ideal gas or a dilute solution, for example, comes a certain entropy \( S_0 \). Let a part of \( V_0 \) be \( V \), and all the \( n \) points be imagined to be moved into \( V \) without any other change in the system. Another value \( S \) of entropy comes to this state. The difference in entropy can be determined with the help of the Boltzmann principle.

Hence, the probability of the second state relative to the first one is to be determined. In other words, one has to know the probability that all the \( n \) points moving independently of each other in the given volume \( V_0 \) are found in the volume \( V \) at a certain time selected by chance. This statistical probability is

\[
W = \left( \frac{V}{V_0} \right)^n. 
\]  

(106)

The application of the Boltzmann principle gives

\[
S - S_0 = R \left( \frac{n}{N} \right) \log \left( \frac{V}{V_0} \right). 
\]  

(107)

It is remarkable that for the derivation of this equation, from which Boyle–Gay–Lussac's law and the law of identical osmotic pressure can be easily derived thermodynamically, one needs to make no assumption about the law of motion of the molecules.

For the entropy of the monochromatic radiation, Eq. (103) is obtained above from Wien's formula. Equation (103) is rewritten as

\[
S - S_0 = \frac{R}{N} \log \left[ \left( \frac{V}{V_0} \right)^{N/\hbar^2 \nu} \right]. 
\]

(108)

From the comparison of this equation with the general equation of the Boltzmann principle, Eq. (105), there derives the following conclusion.

"If monochromatic radiation of the frequency \( \nu \) and energy \( E \) is enclosed in the volume \( V_0 \) (by reflecting walls), the probability that at a certain time selected by chance the total energy of radiation is found in the partial volume
V of the volume $V_0$ is

$$W = \left( \frac{V}{V_0} \right)^{\frac{N}{\kappa} \cdot \frac{\varepsilon}{\beta \nu}}.$$  \hspace{1cm} (109)

Because the comparison of this with the probability given for the case of the ideal gas in Eq. (106) gives $n = (N/R)(E/\beta \nu)$, Einstein arrives further at the following conclusion.

"Monochromatic radiation of low density (within the bounds of validity of Wien's radiation formula) behaves in thermodynamical relations, as if it consisted of mutually independent energy quanta of the quantity $R \beta \nu/N$".

Einstein compares also the average quantity of the energy quanta of black body radiation with the average kinetic energy of the center-of-mass motion of a molecule at the same temperature. The latter is $(3/2)(R/N)T$, whereas for the average quantity of the energy quanta one has from Wien's formula

$$\frac{\int_0^\infty \alpha \nu^3 e^{-\beta \nu/T} d\nu}{\int_0^\infty (N/R \beta \nu) \alpha \nu^3 e^{-\beta \nu/T} d\nu} = \frac{3}{N} \frac{R}{T}. \hspace{1cm} (110)$$

Einstein says that "if monochromatic radiation (of sufficiently low density) behaves now regarding the dependence of the entropy on the volume like a discontinuous medium, which consists of energy quanta of the quantity $R \beta \nu/N$, there arises the examination whether the laws of generation and absorption of light also are made as if light consisted of such energy quanta", and treats Stokes's law and the photoelectric effect.

§8. Stokes' law and the light quantum

Einstein assumes that also in the case of the change of monochromatic light into light of other frequency by photoluminescence, the generated as well as generating light consists of energy quanta of the quantity $(R/N)\beta \nu$, where $\nu$ is the frequency concerned. Every generating energy quantum of the frequency $\nu_1$ is supposed to be absorbed and to give rise to a light quantum of the frequency $\nu_2$. If the photoluminescent substance is not to be regarded as a constant source of energy, the energy of a generated energy quantum cannot be greater than that of a generating light quantum from the principle of conservation of energy. There must hold therefore the relation

$$(R/N)\beta \nu_1 \leq (R/N)\beta \nu_2,$$
from which it results that

\[ \nu_1 \leq \nu_2. \]

This is the well-known Stokes's law. Einstein proves thus from the theory of light quantum this law, which could not be explained from the classical electromagnetic theory.

From this new view it becomes clear that, in the case of weak exposure, the amount of the generated light must be proportional to the intensity of the exciting light, because every exciting energy quantum will cause an elementary process, independently of the action of the other exciting energy quanta. In particular, there will be no limit to the intensity of the exciting light, below which the light would be unable to have exciting effect.

According to the view explained above, it is also clear that Stokes' law is not valid in the following cases:

1. when the number of the energy quanta existing in the process at one time per unit volume is so great, that an energy quantum of the generated light can receive its energy from a number of exciting energy quanta;
2. when the generating (or generated) light is not within the bounds of validity of Wien’s formula.

§9. *Photoelectric effect and the light quantum*

Einstein treats next the photoelectric effect. He says that “the usual view, that the energy of light should be distributed continuously over the illuminated space, finds in attempting to explain photoelectric phenomena particularly great difficulties, which are set forth in a pioneering work of Mr. Lenard”.

According to the view that the exciting light consists of energy quanta of the energy \((R/N)\beta \nu\), the generation of cathode rays by light can be seen as follows. “Into the superficial layer of the body energy quanta penetrate, and their energy is converted at least in part to the kinetic energy of electrons. The simplest idea is that a light quantum hands over its whole energy to a single electron; we want to assume that this should occur. It is, however, not to be excluded, that electrons receive the energy of light quanta only partially. One electron provided with kinetic energy in the inside of the body will, when it has reached the surface, have lost a part of its kinetic energy. It will be, besides,
to be assumed, that every electron has done a work $P$, which is characteristic of the body, when it leaves the body”.

In these words, electrons were of course thought utterly classical particles. Any expression like the one used in later times that an electron makes a quantum transition from a state in the metal to a state of free motion, was not yet arrived at. This kind of idea was later presented for the first time by Bohr,\textsuperscript{42} as is well known. However, Einstein did not already think the change in the motion of electrons in the classical electromagnetic-theoretical way. Each electron was supposed to get suddenly a kinetic energy on absorbing a finite amount of energy from a light quantum. In this respect Einstein’s thought can be said to be the predecessor of the Bohr theory. In other words, Einstein’s understanding was more substantialistic. On this account, he had no need to use particular expressions such as quantum transition, and was able to take a view more by means of model. His substantialistic understanding was made possible also by the fact, that the phenomenon treated was itself the one in which the electron was finally not in a bound stationary state of the atom and therefore could be treated with the classical corpuscular image as it was.

Now, in the paper under consideration, Einstein discusses that the kinetic energy of the leaving electron is given by

$$(R/N)\beta \nu - P.$$  

If the body is charged up to the potential $\Pi$ and surrounded by a conductor of the null potential, and if $\Pi$ is just able to prevent any loss of the electricity of the body, there must holds the relation

$$\Pi e = (R/N)\beta \nu - P,$$

where $e$ is the electric charge of the electron, or the relation

$$\Pi E = R\beta \nu - P',$$

where $E$ is the electric charge per unit gram equivalent of monovalent ions and $P'$ the potential which the negative charge of this quantity has with respect to the body.

Einstein compares the last equation with Lenard’s experiment. If one sets $E = 9.6 \times 10^3$, $\Pi \times 10^{-8}$ is the potential in volts which the body takes in the case of illumination in vacuum. He puts, to see the agreement of his

\textsuperscript{42}N. Bohr, Phil. Mag. 26 (1913), 1.
result with experiment in the order of magnitude, $P' = 0$ and $\nu = 1.03 \times 10^{15}$ (corresponding to the limit of the solar spectrum towards the ultraviolet), getting $\Pi \times 10^7 = 4.3$ volt. This agrees with Lenard’s result$^{43}$ in the order of magnitude.

If the derived formula is right, $\Pi$ must be, on being represented as a function of the frequency of the exciting light in Cartesian coordinates, a line, whose slope is $R\beta\nu$ independently of the nature of the substance investigated.

Einstein’s theory is thus seen not to be in contradiction with the characteristics of the photoelectric effect observed by Lenard. If every energy quantum of the exciting light hands over independently of all the others its energy to electrons, the velocity distribution of the electrons, that is, the quality of the generated cathode rays will be independent of the intensity of the exciting light. On the other hand, the number of the electrons leaving the body will be proportional to the intensity of the exciting light under otherwise the same circumstances. These are in agreement with Lenard’s experiment.

About the supposed limits of validity of the mentioned conformity with facts, remarks similar to those regarding Stokes law would be to be made.

Though Einstein supposes that the case is most probable in which the energy of every light quantum is completely absorbed, he considers also other cases to assume the relation

$$\Pi E + P' \leq R\beta\nu.$$  

For the cathode ray luminescence, which is the inverse process of the photoelectric effect, he assumes the relation

$$\Pi E + P' \geq R\beta\nu,$$

from an analogous consideration. Einstein examines thus Lenard’s experimental results obtained in 1903.$^{44}$ For the substances studied by Lenard $\Pi E$ is always considerably greater than $R\beta\nu$. That is, the voltage which the cathode rays must pass through in order to be able to generate visible light, amounts to several hundred volts in some cases and to several thousand volts in other cases. Einstein says that “it is therefore assumed that the kinetic energy of an electron is used for the generation of many energy quanta of light”. This

$^{43}$P. Lenard, Ann. d. Phys. 8 (1902), 165, 184.

$^{44}$P. Lenard, Ann. d. Phys. 12 (1903), 469.
corresponds to the effect of bremsstrahlung, which became a particular subject of detailed study nearly from that time in relation to the generation of X-rays.

§10. Ionization of molecule

Einstein discusses then the ionization of gases by ultraviolet rays. He considers that in this case a light quantum is absorbed to ionize a gas molecule. From this it follows that the ionization work (that is, the work needed theoretically for the ionization) of a molecule cannot be greater than the energy of an absorbed light quantum. If \( J \) is the theoretical ionization work per gram equivalent, one must have

\[
R\beta \nu \geq J.
\]

According to Lenard's measurements the longest effective wavelength for air is about \( 1.9 \times 10^{-15} \) cm, so that one has

\[
R\beta \nu = 6.4 \times 10^{12} \text{erg} \geq J.
\]

One gets an upper limit for the ionization work from the ionization voltages of dilute gases, for which Stark's work\(^{45}\) can be used. The smallest measured ionization voltage (on platinum anodes) for air is about 10 volts. There results therefore an upper limit \( 9.6 \times 10^{12} \) for \( J \), which is nearly equal to the value found above.

To conclude, Einstein says as follows. "There results still another consequence, of which the test by experiment seems to me to be of great importance. If every absorbed energy quantum of light ionizes a molecule, there must exist between the absorbed quantity \( L \) of light and the number \( j \) of the gas molecules ionized by it the relation

\[
i = \frac{L}{R\beta \nu}.
\]

This relation must, if our view is in accordance with the reality, be valid for every gas, which (at the frequency considered) shows no noticeable absorption unaccompanied by ionization".

\(^{45}\)J. Stark, "Die Elektrizität in Gasen" (Leipzig, 1902), p. 57.
§11. Einstein’s analysis of Planck’s theory

Einstein’s paper under consideration is really a bright one. In this paper he insistently did not accept the universal constant $h$ used by Planck. Einstein used the constant $\beta$ as the one to fit the phenomenon of thermal radiation, without giving it any particular meaning, while he used the combination $(R/N)\beta$ as the constant for light quanta. Of course one sees that

$$
\left(\frac{R}{N}\right)\beta = h.
$$

It can be said that Einstein would not use $h$ because he adopted Planck’s radiation formula as a formula rather close to an empirical one without sufficient foundation, ignoring the fact that Planck discovered the energy quantum.\(^{46}\)

He expressed thus the constant for the energy quantum in terms of the constants of definite origin. This came to be emphasized still more on account of the use of $R\beta$ to fit the experimental facts of the photoelectric effect.

In the next year 1906 Einstein\(^{47}\) came to accept Planck’s quantum theory, as we have mentioned before in §4 of the present section. In this paper Einstein examines, on the basis of the kinematics treated by him in 1903,\(^{48}\) Planck’s idea which is different from Einstein’s view of the light quantum presented in his foregoing paper, that is, the idea which started from the premise that resonators could take only specified values of energy. In other words, Einstein shows that, if one calculates the entropy of a system of resonators by applying as usual the method of Boltzmann’s theory of gases, one arrives at Rayleigh-Jeans’s formula which does not fit the facts, and that, on the other hand, “one arrives at Planck’s formula by assuming that the energy of a resonator cannot take any arbitrary value, but only the values which are integral multiples of $\varepsilon$, where

$$
\varepsilon = \left(\frac{R}{N}\right)\beta\nu^\prime
$$

Einstein points out then that the probability $W$ concerned becomes the quantity which Planck called “the number of complexions”.

\(^{46}\)Cf. §4 of the present section.


Einstein states that "we must hence regard the following law as giving rise to Planck’s theory:

The energy of an elementary resonator can take only the values which are integral multiples of \((R/N)\beta\nu\); the energy of a resonator changes by absorption and emission in a jumping way, and to be precise by an integral multiple of \((R/N)\beta\nu\). In this statement, we can see the first flash of the idea of stationary states and quantum jumps. Moreover, the contradiction with the classical theory emphasized by Bohr is here pointed out in a clear way. Also pointed out is the contradiction in Planck’s theory that both the classical theory and the new idea were used in its development. In other words, this assumption contains the second idea that makes a part of Planck’s theory but is contradictory with the electromagnetic theory, on the basis of which Eq. (89), that is

\[
\overline{E}_\nu = \frac{c^3}{8\pi\nu^2} u_\nu ,
\]

(113)
is derived for the relation of the average energy of a resonator to the energy density of radiation.

Einstein says that "if the energy of a resonator can change only in a jumping way, to the determination of the average energy of a resonator existing in the space of radiation, the usual theory of electricity cannot be applicable”, and points out that the following assumption gives rise to Planck’s theory.

"Although the Maxwell theory is not applicable to elementary resonators, the average energy of an elementary resonator exiting in the space of radiation is nevertheless equal to the one which one calculates by means of the Maxwell theory of electricity”.

This sentence would quite be plausible, if \(\varepsilon = (R/N)\beta\nu\) is small compared to the average energy \(\overline{E}_\nu\) of a resonator. But, “this is not necessarily the case”. In other words, within the bounds of validity of Wien’s radiation formula \(e^{\beta\nu/T}\) is large compared to 1. According to Planck’s theory of radiation \(\overline{E}_\nu/\varepsilon\) has the value \(e^{-\beta\nu/T}\) within the bounds of validity of Wien’s radiation formula, so that \(\overline{E}_\nu\) is far smaller than \(\varepsilon\). A value of energy different from zero comes therefore only to a few resonators.

Einstein arrives at the conclusion that “the above considerations do not refute Planck’s theory of radiation; they seem rather to show that in his theory of radiation Mr. Planck has introduced a hypothetical element — the hypothesis of light quantum — into physics”.
Einstein's indication of the contradiction inherent in Planck's theory has
an important significance, of which we shall later give a full description.\textsuperscript{49)}

It was in his paper in 1909,\textsuperscript{50)} that Einstein presented for the first time
the problem of unifying both the undulatory and corpuscular characters. He
states, in this paper, "For that reason it is my opinion, that the next phase
of the development of theoretical physics will bring us a theory of light, which
can be grasped as a kind of fusion of undulatory and corpuscular theories of
light". We see that the indication is here given of the contradiction about
the substantialistic light quantum, and at the same time of its solution in an
essentialistic direction. He states also that "to give reasons for this opinion,
and show that a profound alteration of our view of the essence and constitution
of light is indispensable, it is the purpose" of his paper. We shall treat this
problem in Vol. II.\textsuperscript{51)}

\section{The Light Quantum and the Theory of Relativity\textsuperscript{52)}}

On the introduction by Einstein of the theory of special relativity, there have
hitherto been presented quite Machist and idealistic interpretations. In other
word, in accordance with the guiding principle of the Machist empiricism, that
physics is not to treat objects but to arrange and describe sensory experiences,
it has been said that the theory of relativity was formulated "phenomenologically"
by the exclusion of the unperceivable and unobservable ether out of
physics.

By practical and careful examinations, however, we can see almost no color
of the Machist empiricism in Einstein's papers and also in his lectures. The
interpretation from the side of empiricism of the theory of relativity has been
made by people other than Einstein. We can rather say that Einstein has been
a realist with a religious color of Planck's sense.

In particular, Einstein had an intense realistic color, in the time until about
1910, when the theory of light quanta had still an intense substantialistic color.

Einstein's theory of light quanta and his theory of relativity appeared in the
same year 1905. Nevertheless, about the relation between both the theories,

\textsuperscript{49)}Translator's note: Cf. Sec. 1, Chap. 1, Vol. II of the present book.
\textsuperscript{51)}Translator's note: Cf. Sec. 3, \S 3, Chap. 1, Vol. II of the present book.
\textsuperscript{52)}Translator's note: Being translated by another translator into English, the present sec-
it is commonly said that they are not related to each other in any way. For example, in de Broglie's book "Matter and Light", it is described that "the first development of the quantum theory (1900 ~ 1916) is effected quite in fact independently of the idea of relativity".53)

It is very interesting to see the sequence of both the theories. Einstein wrote his paper on the theory of light quanta in Mar. 1905, and his first paper on the theory of special relativity "On the Electrodynamics of Moving Bodies" in Jun. of the same year. We should consider that this sequence is not merely accidental.

As a matter of fact, we have reason to say that Einstein abandoned ether not only on account of Michelson-Morley's experiment, but also because of the viewpoint of the theory of light quanta. As we have mentioned before, in the theory of light quanta Einstein conceived an idea somewhat similar to that in Newton's corpuscular theory of light. It is quite different from the idea in those days, namely, the idea of seeing light as an elastic wave in ether, which is incompatible with it. With the idea of light as an oscillation of ether, it is impossible to treat light, as Einstein did, as a group of molecules to which the kinetic theory is applied. Also, Einstein stated already in the former paper that the classical electromagnetic theory and the theory of light quanta are sharply in opposition to each other. In this paper Einstein did not treat light by the idea of light as an oscillation of the ether. Thus, we should say that the negation of the other was not made for the first time in his paper on the theory of relativity, but was made for the first time in his paper on the theory of light quanta.

We could say now that, since the medium named ether was actively negated, there arose the necessity of establishing newly such kinematics as was not grounded on ether. In other words, this did not arise from the passive viewpoint that, as is said in the Machist empiricism, it was abandoned because it was not observed.

That part of ether, which was supposed for the nature of light or electromagnetic waves in matter, was first dissolved by Lorentz's theory of electrons, into more definite substances and their motions in atomism. This removed one

53) L. de Broglie, "Matière et Lumièrè" (Abbin Michel, 1937), p. 229. He says then that "there is however an aspect of the quantum hypothesis, aspect indicated by Einstein, which has a relationship to the relativity". But, these words are concerned with the transformation law of the energy and momentum of light, and are not in the meaning under our present discussion.

of the contradictions about ether, but there still remained the ether at absolute rest for light or electromagnetic waves in vacuum. The latter was then dissolved into a more definite substance of light quantum. The abandonment of ether was not based on the guiding principle of sensualism, but was the solution in a materialistic direction. In the first paper on the theory of relativity, Einstein did not bring out with caution the theory of light quanta as it was. In "§8. Transformation of the Energy of Light Rays. Theory of the Pressure of Radiation Exerted on Perfect Reflectors" of this paper, however, he did not simply take the wave theory for light, using such an expression as "a given light complex".

The relation under consideration between the theory of light quanta and the theory of relativity is most clearly described in Einstein's paper in 1909, "On the Development of Our View of the Essence and the Constitution of Radiation".\(^{55}\) In other words, in this paper he treats the relation between the problem of ether and the light quantum.

Einstein says in this paper — "Today we must, however, consider the hypothesis of ether an overcome viewpoint. It is indeed undeniable, that there is an extensive group of facts concerning radiation, which show that certain fundamental characteristics must be attached to light that can be understood far easily from the viewpoint of Newton's emission theory of light than the viewpoint of the undulatory theory".

Einstein discusses then the change of the hypothesis of ether to treat Michelson-Morley's experiment, and, after having traced the logical process of negating ether, treats the electromagnetic field, to say that "if one refuses the hypothesis of ether \ldots, the electromagnetic fields constituting light do not then appear to be states of a hypothetical medium, but appear to be independent objects (Gebilde), which are emitted from the sources of light, just as according to Newton's emission theory of light", and that "a space not got accepted by radiation and free from ponderable matter appears actually to be empty". These words indicate that the negation of ether and the theory of light quanta are thought to be related to each other.

Einstein next shows on the basis of the theory of relativity that a body diminishes its mass by \(L/c^2\) when it emits radiation of energy \(L\).

He considers a free body at rest having initially the energy \(E_0\), which emits the halves of the radiation energy in two opposite directions to each other, to

have the energy $E_1$. From the principle of energy it follows that

$$E_0 = E_1 + L.$$  \hspace{1cm} (114)  

He observes now this from the coordinate system, relatively to which the body moves with the velocity $u$. If every quantity in this system is primed, from the theory of relativity it results for the radiation that

$$L' = L \frac{1}{\sqrt{1 - u^2/c^2}}.$$  \hspace{1cm} (115)  

Because the principle of energy must hold concerning this system, it is found that

$$E'_0 = E'_1 + L \frac{1}{\sqrt{1 - u^2/c^2}}.$$  \hspace{1cm} (116)  

By expanding this in powers of $u/c$ and neglecting the terms of orders higher than $(u/c)^2$, he gets from the difference between the result and Eq. (114)

$$(E'_0 - E_0) = (E'_1 - E_1) + \frac{1}{2} \frac{L}{c^2} u^2.$$  \hspace{1cm} (117)  

$E'_0 - E_0$ is nothing but the kinetic energy of the body seen from the moving system before the emission of light, and $E'_1 - E_1$ is nothing but that after the emission of light. If the mass of the body before the emission is denoted by $M_0$ and that after the emission by $M_1$, there results the relation

$$\frac{1}{2} M_0 u^2 = \frac{1}{2} M_1 u^2 + \frac{1}{2} \frac{L}{c^2} u^2,$$  \hspace{1cm} (118)  

or

$$M_0 = M_1 + \frac{L}{c^2}.$$  \hspace{1cm} (119)  

About this Einstein says as follows — “The theory of relativity has therefore changed our view of the nature of light insofar as light is not grasped as a result of states of a hypothetical medium, but as something independently existing like matter. In addition, according to this theory there is the characteristic in common with any corpuscular theory of light, of transferring inertial mass from the emitting to the absorbing body”.

After having thus explained from the theory of relativity the necessity of considering light like a material particle, Einstein makes clear then that from the undulatory theory no explanation can be deduced of the photochemical
reaction, photoelectric effect, etc. He says that "it seems that in connection with this point the emission theory of light by Newton preserved more truth than the undulatory theory, since according to the former the energy, which is given to a particle of light in the emission, is not scattered over the infinite space, but remains available for an elementary process of absorption", and enters into an explanation of the theory of light quanta. The theory of relativity and the theory of light quanta made their starts in such a close relationship with each other as we have described in the present section.\(^{56}\)

\(^{56}\)Translator’s note: Cf. on this point the middle part of Sec. (2) of Explanatory Notes added to the republished edition in 1972 of this Vol. I by means of photocomposition, which are given at the end of the present volume.
Chapter 2
The Formation of Atomic Models

1. On Line Spectra

§1. *Experiments on line spectra*

Continuous spectra emitted by solids at high temperatures gave birth to the quantum theory, as we have described in the last section, by virtue of the theory of molecular linear oscillators, which started from the theory of black body radiation. On the other hand, line spectra were made clear by the theory of atomic structure. So they remained being considered only in phenomenological descriptions of experiment, until the substantialistic knowledge of atomic structure became established.

It was T. Melvill (1726–1753) in England who found for the first time spectral lines coming out of shining gases. He dropped rock salt, salt, potassium sulfate, niter and so on into a flame of alcohol, and put a sheet of pasteboard with a small hole before the flame, placing a prism at the back of the sheet. He found in this way light of somewhat various colors depending on matter put into the flame, and saw that yellow light was shining bright in particular remarkably over light of the other colors. This means the $D$-line of natrium. He found also that the change of color from this shining yellow to the other weak ones was sudden and not gradual. W. H. Wallaston (1766–1828) in England found next a band of blue light at the bottom of the flame of a candle. This is the spectrum called Swan’s spectrum nowadays, of which W. Swan gave a description in 1856. Of the dark lines in the solar spectrum seven were first found by Wallaston in 1802. He considered that they were borderlines between purely monochromatic parts of the spectrum.
Epoch-making achievements were performed about line spectra by J. Fraunhofer (1787–1829) in Bavaria. He made, independently of Wallaston's study, the first great study of solar dark lines. Because he had grown in the home of a poor glass polisher, he became skilled in glass polishing by virtue of helping his father. This played an important role in his researches in optics. He made for the first time the achromatic lens. Studying the refractive index of glass for specified colors to make more effective achromatic lenses, he found an orange double line in the light of a lamp. This was the double line of natrium. He recognized the sharp and clear double line in the flames of oil, fat and all the other combustible substances. He wrote that it was very useful to the determination of refractive index, because it appeared correctly at one and the same position.

He separated the light passed through a fine slit by a prism of flint glass placed at a long distance from the slit, and observed it by the telescope of a theodolite. He took also the sunlight as the source, and tried to find bright lines in it as in the light of a lamp. However, he found instead of detecting such ones very many strong or weak vertical lines, which were darker than the other parts and almost completely black. He found bright lines in the case of hydrogen, alcohol and sulfur. This was because of the contamination by natrium. He examined starlight and found in the case of Venus the same lines as those of the sunlight.

Fraunhofer was the first who made observation with the use of (diffraction) grating. By means of it he measured for the first time the wavelength of the D-line in 1821. He got the value of 0.0005888 mm, which should be said very good.

It is said,\footnote{F. Cajori, "A History of Physics" (1929), p. 165.} that Fraunhofer's works were little noticed, because at that time optics was in the midst of the controversy between the emission and undulatory theories, and chemistry concentrated on Darton's theory. Fraunhofer himself could not solve the mystery of the solar dark lines, that is, of the Fraunhofer lines, and could not find that spectral lines play an important role in chemical analysis.

In 1827 J. F. W. Herschel in England stated in his book "On Light", about the result of his examination of several substances, that a small amount of matter could be detected by means of bright lines of various colors. In 1835,
C. Wheatstone in England gave some descriptions of the spectra of electric arcs between metals.

Kirchhoff criticized that English researchers did not set up the relation between a specific element in a flame and spectral lines. In fact, W. H. F. Talbot (1800–1877), for example, attributed the D-line both to sulfur and natrium.

In 1832 David Brewster made clear that spectral black lines arose from the absorption of light when it passed through a colored glass or specific glass. He found that these black lines were similar to those in the spectrum of the sunlight. He found also that, while vapouring nitric acid presented absorption lines, in the case of liquid one no absorption took place, and for this reason denied the undulatory theory of light.

W. A. Miller in England and J. B. L. Foucault at Paris made clear that the bright line of natrium and the black D-line of the sunlight were completely coincident with each other. Foucault proved this by leading the sunlight and the light from an electric lamp presenting the natrium line into a spectroscope at the same time.

J. N. Niepce’s invention of photographic technology in 1827 had a great meaning to the research of spectra. J. W. Draper at New York applied to his research L. J. M. Dauguerre’s invention of daguerreotype in 1839, and took a photograph of Fraunhofer lines in 1842. Draper made an excellent study on the thermal radiation from solid, showing in 1847 that all substances emit visible light at temperatures greater than 525°C. He deduced from his study a wrong conclusion that liquid as well as solid had a continuous spectrum, and it was accompanied by a line spectrum. This was caused by the mixture of the line spectrum of a metallic acid in the flame used, and by that of the continuous spectrum of solid carbon.

There were also excellent achievements by A. Angström, E. Bequerel, and others, and in particular there appeared in 1859 epoch-making works by R. W. Bunsen and by Kirchhoff, which brought spectroscopy on her definite course. These great physicist and chemist made close collaboration in the university of Heiderberg, and laid the foundation of spectral analysis in the time 1859–1862.

---

In 1855 gas for illumination was laid on in laboratories, and in 1857 the Bunsen burner was thought out by Bunsen and Roscoe. The role played by the Bunsen burner in spectral analysis was great, because it supplied flames of high enough a temperature but with no shine. Chemical substances were vaporized by these flames to become purely simple gases capable of shining and giving their characteristic spectra. Most serious parts of the mistake in former experiments caused by the disturbance of the spectrum of flame were thus avoided.

Kirchhoff gave in his studies important elucidation of the black lines of the solar spectrum. A flame having a line spectrum weakens light of the colors corresponding to its bright lines. In other words, he showed that the bright lines of a flame turn into black lines, if bright enough light is put behind the flame, and therefore the solar black lines are not due to the absorption by the air of the earth, but due to substances existing in the shining solar atmosphere. Kirchhoff concluded thus that Na, Fe, Mg, Cu, Zn, Ba and Ni exist in the solar atmosphere.

This phenomenon of spectral conversion laid the important foundation of Kirchhoff's law of radiation as we have mentioned previously. The law established by Bunsen and Kirchhoff stated that every set of bright lines was to mean a definite proof of the existence of a corresponding metal. This law got its certainty and effectiveness still more through the discovery of two new metallic elements in mineral water from Durkheim by spectral analysis. One of them was distinguished by blue lines and the other by red lines. They were named by these scientists cesium and rubidium, respectively.

The application of spectral analysis to astronomy was done mainly by Kirchhoff, and his explanation of the Fraunhofer lines was epoch-making.

Bunsen-Kirchhoff's law was to give the correspondence of a specific line spectrum with a specific element, and was thus a phenomenological law. It was, however, a systematization of the former results in spectroscopy, and was the first experimental law possessed of the nature of scientific investigation in this field.

After Bunsen-Kirchhoff's work active experimental researches were made in the field of spectroscopy. In particular, researches were advanced by means of the Geissler tube (Plücker tube), which was designed by J. Plücker at Bonn in 1858 and was made by H. Geissler at Plücker's request. In 1862 Plücker showed that one and the same matter had different spectra at different temperatures.
Plücker and W. Hittorf found that each vapour of hydrogen, nitrogen and sulfur had a weak band spectrum as well as a bright line spectrum.

A. Wüllner discussed that hydrogen, oxygen and nitrogen changed their spectra by a change in pressure of a Geissler tube. He obtained three kinds of spectra for oxygen by changing the pressure, and thought that both the pressure and temperature brought about a change in the spectrum. Wüllner's conclusion was criticized by Ångström, who attributed some change to impurities and pointed out that, though a change in the temperature caused a change in the spectrum, an increase in the pressure did not give rise to a new spectrum, only broadening the lines in width.

In 1873–1874, J. N. Lockyer asserted that a line spectrum was due to a free atom, and a band spectrum due to a molecule. Ångström was in favor of this, but Wüllner made an experiment on nitrogen in 1879, and showed that the band spectrum shifted gradually to a line spectrum as the temperature was gradually changed. He thought that Lockyer's theory of molecular dissociation was not necessary to explain the fact. Lockyer found that the line spectrum changed with the temperature, in the case of calcium, for example. He advanced his former theory of considering the shift of the band spectrum to the line spectrum due to the dissociation of a molecule into atoms, and supposed that an atom would dissociate into more elementary substances, assuming chemical elements to be composite.

In 1884 J. J. Balmer discovered a mathematical relation among the spectral lines of hydrogen, and showed that they were correlated to form a series.

From 1890 important researches were started by H. Kayser and C. Runge. They showed about various elements that the distribution of spectral lines was not so random as it had formerly been considered. In other words, they found that the spectral lines of each element formed series, as was shown by Balmer for hydrogen. In the case of argon, they supposed the mixture of other elements, because more than one series were found. This supposition when applied to oxygen which had six series, ended in complete failure, and was abandoned.

In 1896 P. Zeeman made elements to shine in a magnetic field, and observed changes in their spectra. Faraday had tried this kind of experiment already in 1862, but had observed no change. Faraday had put a flame between magnetic poles. Zeeman made the observation possible with the use of an improved device, and found that the lines broadened in their widths. A. A. Michelson
at Chicago found by means of an echelon spectroscope that this effect was more complicated.

§2. Mathematical law of line spectra

In this way the experimental study of line spectra made progress from the beginning of the 19th century, and established the correspondence of the existence of each element with its characteristic spectral lines. This was a phenomenological knowledge, but was the first one possessed of a substantialistic character in the sense that it meant the existence of elements. Over visible and near ultraviolet spectra, and of many kinds of gases and vapour, investigations were carried out and the wavelengths of spectral lines were tabulated. It was in 1884, however, that some lawfulness was discovered of the distribution of spectral lines for the first time.

This discovery was done by J. J. Balmer (1825–1898), who was a teacher at a middle school and an unofficial lecturer at the University of Basel. It was at an age of his 60's. Balmer found out, about the four visible lines Hα, Hβ, Hγ, Hδ and several lines in the ultraviolet region of the hydrogen spectrum, that their wavelengths were in a simple mathematical relation. The series of these lines is called nowadays the Balmer series, and is expressed by the following formula. That is, according to the usage in spectroscopy, the wave number (the number of waves per cm) of each line $\nu = 1/\lambda$, $\lambda$ being the wavelength, is given by

$$\nu = \frac{1}{\lambda} = R \left( \frac{1}{m^2} - \frac{1}{5^2} \right) \text{ cm}^{-1}, \quad (m = 3, 4, 5, \ldots),$$

where $R$ is called the Rydberg constant equal to

$$R = 109677 \cdot 691 \text{ cm}^{-1},$$

and $m = 3, 4, 5, 6$ give Hα, Hβ, Hγ, Hδ, respectively.

In his paper under consideration, Balmer states as follows — "Ångström's very precise measurement of the four hydrogen lines made it possible to call on a common factor for their wavelengths, which stood by the wavelengths in as far as possible simple numerical relationship. Thus I arrived gradually at a formula, that can be regarded at least for these four lines as the expression of a law, by which their wavelengths are represented with surprising accuracy."

---

For the common coefficient in this formula, Balmer obtains from Ångström’s values of the wavelengths the value

\[ h = 3645.6 \text{ mm}/10^7, \]

\( h \) not denoting here the Planck constant. About this constant Balmer gives a very foresighted opinion, that “one could name this number the fundamental number of hydrogen, and if one should succeed in finding for other elements the corresponding fundamental numbers of their spectral lines, the supposition would be permitted, that between their fundamental numbers and corresponding atomic numbers, determinate and also by some function expressible relations take place”.

The first four wavelengths are obtained by multiplying respectively the coefficients

\[ \frac{9}{5}, \frac{4}{3}, \frac{25}{21}, \frac{9}{8} \]

by the fundamental number. Though at first sight these coefficients seem to be a sequence without any regularity, the multiplication by 4 of both the denominator and numerator of the second and fourth of them makes immediately their regularity clear. The coefficients so obtained are found to have the numbers

\[ 3^2, 4^2, 5^2, 6^2, \]

as their respective numerators, and the numbers respectively equal to them minus 4 as the denominators.

Balmer induces by generalizing this the following — “It is likely to from various reasons, that to the four coefficients given above two series belong, so that the second series takes in the terms of the first series again.” Thus he obtains the formula for the coefficients in general

\[ \frac{m^2}{m^2 - n^2}, \]

where \( m \) and \( n \) are integers. Hence the formula for the wavelengths becomes

\[ \lambda = h \cdot \frac{m^2}{m^2 - n^2}. \]
He calculates the following values from this formula.\(^4\)

<table>
<thead>
<tr>
<th></th>
<th>Calculated</th>
<th>Ångström's measurement</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H_\alpha) (C-line)</td>
<td>((9/5)h = 6562.08)</td>
<td>6562.10</td>
<td>+0.02</td>
</tr>
<tr>
<td>(H_\beta) (F-line)</td>
<td>((4/3)h = 4860.8)</td>
<td>4860.74</td>
<td>-0.06</td>
</tr>
<tr>
<td>(H_\gamma) (pre G)</td>
<td>((25/21)h = 4340)</td>
<td>4340.1</td>
<td>+0.1</td>
</tr>
<tr>
<td>(H_\delta) (h-line)</td>
<td>((9/8)h = 4101.3)</td>
<td>4101.2</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

The relative deviation of this formula from Ångström's measured values does not amount at most to 1/40,000, being within the experimental error. From this Balmer praises conversely Ångström, saying that this “is rather an excellent evidence for the great conscientiousness and care, with which he must have proceeded in his operation”.

With this formula Balmer calculated as the fifth line of hydrogen

\[
(49/45) \cdot 3645.6 = 3969.65 \times 10^{-7} \text{ mm}.
\]

This line would be in the visible region close before the line \(H_1\), that is, the line measured by Ångström to be 3968.1, but it was not known to Balmer. So, he was led to consider that either in former experiments this line was not in a favorable condition to appear on account of the relation to the temperature, or this formula had not general validity.

Balmer was informed by Hagenbach, however, the fact that in the violet and ultraviolet regions measured by Vogel of the spectrum of hydrogen, and also in those measured by Huggins of the spectrum of a white star, there were many hydrogen lines, and was given the result of a comparison of the measured wavelengths with Balmer's formula.

The formula showed a surprising agreement in these regions. Balmer compared it with the results obtained in the visible region by various experimentalists. The differences among their results were rather remarkable. Their results gave in general smaller values than the present-day values.

In the paper under consideration Balmer calculates up to the fifteenth line, and considers that one has to wait for future studies to see whether the spectrum of a white star agrees with or gradually deviates from his result.

\(^4\)The present-day values are as follows. \(H_\alpha\): 6562.79, \(H_\beta\): 4861.33, \(H_\gamma\): 4340.47, \(H_\delta\): 4101.74. Ångström's values were smaller in general.
Balmer gives then a correct outlook for the problem of other elements — "Should the above formula apply only to the single chemical element of hydrogen and not be again found also in the spectral lines of other simple matter with a fundamental number characteristic to this matter? If not, one would have perhaps to assume that the formula entitled to hydrogen is a particular case of a general formula, which on certain conditions turns just into the formula for the hydrogen lines."

Balmer generalized his formula by means of \( m \) and \( n \). It should be noted that in this paper he states the following expectation of the values of \( \lambda \) for \( n \) other than 2 — "The hydrogen lines, which correspond to the formula for \( n = 3, 4, \) etc., and which one can describe as the lines of the third, fourth, etc. orders, are not found in the spectra known up to now; they must develop in a quite new relationship to temperature and density."

Though Balmer was mistaken a little in generalizing his formula with respect to \( n \), the above expectation was correct. In other words, if his original formula

\[
\lambda = h \cdot \frac{m^2}{m^2 - 2^2},
\]

is expressed in the wave number, it becomes

\[
\nu = \frac{2^2}{h} \left( \frac{1}{2^2} - \frac{1}{m^2} \right).
\]

Though Balmer’s generalization with respect to \( n \) was

\[
\nu = \frac{n^2}{h} \left( \frac{1}{n^2} - \frac{1}{m^2} \right)
\]

the actual formula is

\[
\nu = \frac{4}{h} \left( \frac{1}{n^2} - \frac{1}{m^2} \right), \quad (3)
\]

where

\[
R = \frac{4}{h} \quad (4)
\]

is the Rydberg constant. The lines with \( n = 1, 3, 4, \ldots \) are found afterwards in the spectrum of hydrogen. They are as follows.
The Formation of Atomic Models

- $n = 1$ Lyman series (1906) extreme ultraviolet
- $n = 3$ Ritz-Passen series (1908) infrared
- $n = 4$ Brackett series (1922) infrared
- $n = 5$ Pfund series (1924) infrared.

Balmer expects also in this paper that, if his formula is correct for $n = 2$, the spectral lines will accumulate in the direction of the ultraviolet side toward the wavelength 3645.6, but will not go beyond it. He asserts that in the direction of the red side $H_\alpha$ (C-line) is the outermost line, and predicts correctly that "only when lines of still higher orders occur, still further lines arise also in the direction of the infrared side".

An experimental formula for the distribution of spectral lines of other elements was first published in 1890 by J. R. Rydberg.\(^5\) In these papers he makes use of the wave number

$$n = 10^8 \lambda^{-1}, \quad (5)$$

where $\lambda$ is the wavelength measured in Å, that is, in $10^{-8}$ cm, in order to get better insight as we have mentioned above.

Rydberg treats for the first time the first three families in the periodic table, and induces the laws of doublet and triplet terms. He shows that the respective lines of each series are given by the formula

$$n = n_0 - \frac{N_0}{(m + \mu)^2}, \quad (6)$$

where $m$ is a positive integer denoting the term number. He gives $N_0 = 10921.6$, and considers it a number common to every series and every element, which is the constant named the Rydberg number now. $n_0$ and $\mu$ are constants characteristic of each series, and $n_0$ is the limiting value to which $n$ tends as $m \to \infty$.

About the doublet and triplet terms Rydberg mentions that their existence was pointed out for the first time by Mrs. Liveing and Dewer. Rydberg arranges these series, analyzing them into three series of the diffuse series (the first subordinate series), the sharp series (the second subordinate series) and the principal series. He finds four different series for the families I and III, and six for the family II.

\(^5\)J. R. Rydberg, Phil. Mag. (5) 29 (1890), 331; Compt. Rend. 110 (1890), 394; Wied. Ann. 50 (1893), 625.
Rydberg writes Eq. (6) in a more symmetrical form as

\[ \pm \frac{n}{N_0} = \frac{1}{(m_1 + \mu_1)^2} - \frac{1}{(m_2 + \mu_2)^2} , \]  

(7)

and reproduces the principal or sharp series by keeping \( \mu_1 \) and \( \mu_2 \) invariant and varying respectively one or the other of \( m_1 \) and \( m_2 \). Here we see an early presentation of the combination rule.

Since it is already clear that every element has several kinds of series, Rydberg opposes Lockyer’s theory in which the molecular dissociation is supposed to give rise to various kinds of spectra. According to Rydberg, Lockyer’s theory of the dissociation of elements does not agree with the result of Rydberg’s investigation, and Lockyer’s observation of the spectra of Na and K shows only that for shining atoms one should also take into account a situation like the one in the case of sounding bodies where the relative intensities of overtones change with various conditions. Rydberg considers that various kinds of spectra belong to one and the same system.

At nearly the same time, H. Kayser and C. Runge\(^6\) started energetic investigations of spectroscopy. In these works, they express Balmer’s formula as

\[ \lambda^{-1} = A + Bm^{-2} , \]  

(8)

where \( A \) and \( B \) are constants, and, adding one more term, propose the formula

\[ \lambda^{-1} = A + Bm^{-2} + Cm^{-4} , \]  

(9)

to examine the experimental results in detail.

About this point Rydberg considers that Kayser–Runge’s formula as well as his former one is not correct, and proposes the following formula

\[ 10^8 \lambda^{-1} = n = a - b(m + \mu)^{-2} - C(m + \mu)^{-4} . \]  

(10)

In this way spectral lines were given their description, and the mathematical laws relating them were obtained as experimental formulae. These laws described simply the empirical phenomena, but were not such ones as to talk about what the objects were. In other words, these laws were those that should be said to be in a phenomenological stage. Because of their being mathematical descriptions, they would be Pythagorean if they remained staying there. From the viewpoint of Mach’s or Cassirer’s epistemology, physics should get at

\(^6\)H. Kayser und C. Runge, Wied. Ann. 41 (1890), 302; 43 (1891), 385; 46 (1892), 225; 48 (1893), 126.
the end as soon as they were obtained. On the contrary, physics did not make any end, but did make a start. One sees that the Pythagorean and Platonic way of thinking fixeds physics at a certain early stage of development, and considereds that way as if it were everything to the method of physics.

We note, however, that the stage of spectroscopy up to the time under consideration contained substantialistic elements, as long as specific spectra were considered in the correspondence with their respective specific elements, or specific atoms.

§3. *Attempts of deriving spectral laws from dynamical models*

In the later half of the 19th century, while such conscious empirical methods were adopted on one hand by some physicists, that brought the phenomenological way of not considering the structure of objects as in thermodynamics to a creed, attempts were made on the other hand by other physicists, such as Kelvin in particular, of making clear electromagnetic phenomena inclusive of that of ether, on the basis of the Newton mechanics by supposing appropriate dynamical models. In those cases, dynamical models were not considered to reflect by themselves the structure of objects faithfully, differently from the present-day atomic model which is considered to do so, but were supposed only for the sake of getting laws of mathematically the same form. They were anyhow irrelevant to the question whether their structure was the same as that of their respective objects or not.

Fitzgerald\(^7\) treated for the first time a dynamical model according to the theory of vibration, which would give rise to oscillations obeying the same law as the spectral laws of Balmer and others. Rayleigh\(^8\) treated later the problem in a more comprehensive way by means of the theory of vibration in which he was an expert, and it was made clear that the spectral laws were different from those deduced from the theory of vibration.

The model taken in Rayleigh's paper is on the propagation of waves in a combined system of similar bodies. The bodies are supposed to be similarly disposed at equal intervals of \(a\) along a straight line. The position of each body, as displaced from equilibrium, is assumed to be given by a single coordinate, which will be denoted by \(\varphi_r\) for the \(r\)th body. A wave propagating in one

---


\(^8\)Lord Rayleigh, Phil. Mag. 44 (1897), 356.
direction is represented by taking $\varphi_r$ to be proportional to $e^{i(nt+r\beta)}$. Namely, if we see the system at an instance, the disturbance is periodic when $r\beta$ increases by $2\pi$. If the origin is taken to be at $r = 0$, the distance from the origin to the $r$th body is given by $x = ra$, so that one has

$$\varphi_r \sim e^{i(nt+x\beta/a)}.$$ 

This function is periodic for the change in $x = ra$ by $2\pi a/\beta$, which is the wavelength $\lambda$. If $k = 2\pi/\lambda$, one sees that $k = \beta/a$ and that the propagation velocity $V$ is given by $V = n/k$. The aim of the investigation is to find the relation between $n$ or $V$ and $\lambda$. It is in general a principal problem in the theory of waves to find this relation.

Rayleigh assumes that the force acting each body is due solely to the neighbors situated within a limited distance. These forces determine the vibration of the system about its configuration of equilibrium. The simplest case is that, in which there is no interaction between the bodies. The kinetic and potential energies of the system are given in this case by

$$T = \frac{1}{2} A_0 \sum \varphi_r^2, \quad P = \frac{1}{2} C_0 \sum \varphi_r^2,$$

where $A_0$ and $C_0$ are put to be the same for all $r$ from the similarity of the bodies. Each body vibrates independently according to the equation

$$n^2 = \frac{C_0}{A_0}.$$

The frequency is independent of the wavelength in this case. $n$ is independent of $k$, and $V$ given by $n/k$ is inversely proportional to $k$ and directly to $n$. Rayleigh mentions that the propagation of waves in a system of this kind has been considered by Reynolds.

In the general case, $P$ includes products of $\varphi_r$ with the neighbouring coordinates $\cdots \varphi_{r-2}, \varphi_{r-1}, \varphi_{r+1}, \varphi_{r+2}, \cdots$, and a similar statement holds good for $T$. Writing down explicitly only the terms which involve $r$, one has

$$T = \cdots + \frac{1}{2} A_0 \varphi_r^2 - A_1 \varphi_r \varphi_{r-1} - A_1 \varphi_r \varphi_{r+1}$$

$$- A_2 \varphi_r \varphi_{r-2} - A_2 \varphi_r \varphi_{r+2} - \cdots,$$
where $A_1, A_2, \ldots, C_1, C_2, \ldots$ are constants, which are finite for a certain number of terms and then vanishing. The equation for $\varphi_r$ is therefore

$$A_0 \ddot{\varphi}_r - A_1 \dot{\varphi}_{r-1} - A_1 \dot{\varphi}_{r+1} - A_2 \ddot{\varphi}_{r-2} - A_2 \ddot{\varphi}_{r+2} - \cdots$$

$$+ C_0 \varphi_r - C_1 \varphi_{r-1} - C_1 \varphi_{r+1} - C_2 \varphi_{r-2} - C_2 \varphi_{r+2} - \cdots = 0. \quad (15)$$

Since all the $\varphi_r$'s are proportional to $e^{int}$, the double differentiation is accounted for by the factor $-n^2$. Because $\varphi_r$ is also proportional to $e^{i\beta}$, the above equation becomes

$$n^2 (A_0 - A_1 e^{-i\beta} - A_1 e^{i\beta} - A_2 e^{-2i\beta} - A_2 e^{2i\beta} - \cdots)$$

$$= C_0 - C_1 e^{-i\beta} - C_1 e^{i\beta} - C_2 e^{-2i\beta} - C_2 e^{2i\beta} \cdots,$$

or

$$n^2 = \frac{C_0 - 2C_1 \cos ka - 2C_2 \cos 2ka - \cdots}{A_0 - 2A_1 \cos ka - 2A_2 \cos 2ka - \cdots}, \quad (16)$$

in which $\beta$ is replaced with its equivalent $ka$. By this formula $n$ is expressed as a function of $k$ and of the fundamental constants of the system.

In most of the practically important cases, $A_1, A_2, \ldots$ vanish, so that $T$ has the same simple form as in Eq. (11). If $A_0$ is assumed to be unity for simplicity, Eq. (16) becomes

$$n^2 = C_0 - 2C_1 \cos ka - 2C_2 \cos 2ka - \cdots. \quad (17)$$

When the waves are very long, $k$ tends to zero, and one has

$$n^2 = C_0 - 2C_1 - 2C_2 - \cdots.$$

If this limiting value is denoted by $C$, Eq. (17) is rewritten as

$$n^2 = C + 4C_1 \sin^2 \left(\frac{1}{2}ka\right) + 4C_2 \sin^2(ka) + \cdots. \quad (18)$$
An important class of cases is that the frequency diminishes without limit as \( \lambda \) increases. This is the case in which

\[ C = 0. \]

As we have described in Section 2, in the spectral laws the wave number and therefore the frequency appear in the first power instead of the second power. It is thus necessary to consider the case in which Eq. (17) or (18) becomes the first order of \( n \). This is the case, for example, in which only one of \( C_1, C_2, \ldots \) is finite. If \( C_1 \) alone is finite, one gets

\[ n = 2C_1^{1/2} \sin^2 \left( \frac{1}{2} ka \right). \tag{19} \]

A simple case included under Eq. (19) is that of a stretched string without mass by itself, but carrying unit loads at equal intervals of \( a \). In this case \( P \) is given by

\[ P = \cdots + \frac{T_1}{2a} (\varphi_r - \varphi_{r-1})^2 + \frac{T_1}{2a} (\varphi_{r+1} - \varphi_r)^2 + \cdots, \]

where \( T \) is the tension. From Eq. (14b) one then has

\[ C_0 = 2T_1/a, \quad C_1 = T_1/a, \quad C_2 = 0\cdots, \]

which gives \( n \) the same form as Eq. (19). Rayleigh cites Kelvin's wave model\(^9\) which has the same form of \( P \) as the above one.

Another example in which again \( C_2, C_3, \ldots \) vanish is that proposed by Fitzgerald.\(^{10}\) It is the case of a linear system of rotating magnets, which are with their poles close to one another and are given small displacements to small compared with the distance between two adjacent poles. Each restoring force is proportional to the sum of the angular displacements, \( \varphi_r \)'s, of two contiguous magnets, so that \( P \) is given by

\[
\begin{array}{cccccc}
N \cdot & S & N & S & N & \cdots \\
\end{array}
\]

\[ P \propto \cdots + (\varphi_r + \varphi_{r-1})^2 + (\varphi_r + \varphi_{r+1})^2 + \cdots. \]


\(^{10}\)G. P. Fitzgerald, the same as Ref. 7.
In this case, one has $C_1 = -C_0/2$ and therefore

$$n = n_0 \cos \left( \frac{1}{2} k a \right),$$

where $n_0$ represents the value of $n$ appropriate to $k = 0$, i.e. to infinitely long wavelength. This formula gives $n = 0$ for $\lambda = 2a$, in which case $\varphi_{r+1} = -\varphi_r$.

As an example in which $C_1$ and $C_2$ are finite, Rayleigh considers a system of masses each of which is connected to its immediate neighbours on the two sides by an elastic body capable of bending but without inertia. In this case, $P$ is given by

$$P = \cdots + \frac{1}{2} c (2\varphi_{r-1} - \varphi_r - \varphi_{r+1})^2 + \frac{1}{2} c (2\varphi_r - \varphi_{r-1} - \varphi_{r+1})^2 + \frac{1}{2} c (2\varphi_{r+1} - \varphi_r - \varphi_{r+2})^2 + \cdots.$$  

Comparing this with Eq. (146b), one gets

$$C_0 = 6c, \quad C_1 = 4c, \quad C_2 = -c,$$

and then according to Eq. (18)

$$n = 4c^{1/2} \sin^2 \left( \frac{1}{2} k a \right).$$

Thus far the propagation of waves along an unlimited series of bodies has been considered. If it is supposed that the total number of bodies is $m$ and they form a closed chain, $\varphi$ must be such that

$$\varphi_{r+m} = \varphi_r.$$

From this it follows that

$$\beta = k a = 2s\pi/m,$$

$s$ being an integer. Thus, in this case Eq. (17) becomes

$$n^2 = C_0 - 2C_1 \cos(2s\pi/m) - 2C_2 \cos(4s\pi/m) - \cdots.$$  

When the chain, composed of a limited series of bodies, is not closed but open at the ends, the general problem becomes more complicated. A simple
example is that treated by Lagrange, of a stretched massless string, carrying a finite number of loads and fixed at its extremities. The open chain of \( m \) magnets, for which

\[
P = \frac{1}{2} (\varphi_1 + \varphi_2)^2 + \frac{1}{2} (\varphi_2 + \varphi_3)^2 + \cdots + \frac{1}{2} (\varphi_{m-1} + \varphi_m)^2,
\]
is considered by Fitzgerald. The equations are

\[
\begin{cases}
\varphi_1(1 - n^2) + \varphi_2 = 0, \\
\varphi_1 + \varphi_2(2 - n^2) + \varphi_3 = 0, \\
\cdots \\
\varphi_{r-1} + \varphi_r(2 - n^2) + \varphi_{r+1} = 0, \\
\cdots \\
\varphi_{m-2} + \varphi_{m-1}(2 - n^2) + \varphi_m = 0, \\
\varphi_{m-1} + \varphi_m(1 - n^2) = 0,
\end{cases}
\quad (20)
\]
of which the first and last may be brought to the same form as the others, if one introduces \( \varphi_0 \) and \( \varphi_{m+1} \), such that

\[
\varphi_0 + \varphi_1 = 0, \quad \varphi_m + \varphi_{m+1} = 0. \quad (21)
\]
If one assumes

\[
\varphi_r = \cos nt \cdot \sin \left( r\beta - \frac{1}{2}\beta \right), \quad (22)
\]
the first of Eq. (21) is satisfied. The second is also satisfied provided that

\[
\sin m\beta = 0, \quad \text{or} \quad \beta = s\pi/m. \quad (23)
\]
Equation (20) is satisfied if

\[
2 \cos \beta + 2 - n^2 = 0,
\]
that is, if

\[
n = 2 \cos \frac{s\pi}{2m}. \quad (24)
\]
In Eqs. (23) and (24), \( s \) may assume the \( m \) values from 1 to \( m \). In the case of \( s = m \), one has \( n = 0 \) and \( \beta = \pi \), and then from Eq. (22)

\[
\varphi_r = (-1)^r \cos nt.
\]

The equal amplitudes and opposite phases of the contiguous coordinates, i.e. the angular displacements of contiguous magnets, gives rise to zero potential energy, and therefore to zero frequency of vibration. In the case of \( s = 1 \), the angular deflexions are all in the same direction, and the frequency is the highest admissible. If at the same time \( m \) may be very great, \( n \) reaches its maximum value, corresponding to parallel positions of all the magnets. If this value is denoted with \( N \), the generalized form of Eq. (24), applicable to all masses and degrees of magnetization, may be written

\[
n = N \cos \frac{s\pi}{2m}.
\]  

(25)

If \( m \) is great and \( s \) relatively small, Eq. (25) becomes approximately

\[
n = N \left( 1 - \frac{s^2\pi^2}{8m^2} \right).
\]  

(26)

About this result Rayleigh says that “as \( s \) diminishes we have a series of frequencies approaching \( N \) as an upper limit”, and “we are reminded (as Fitzgerald remarks) of certain groups of spectral lines”. According to Rayleigh “a nearer approach to the remarkable laws of Balmer for hydrogen and of Kayser and Runge for alkalis is arrived at by supposing \( s \) constant while \( m \) varies”. This is to mean that Balmer’s formula is given by

\[
n = N(1 - 4m^{-2}), \quad m = 3, 4, 5, \ldots
\]

as we have described before. Rayleigh adds that “in this case, instead of supposing that the whole series of lines correspond to various modes of one highly compound system, we attribute each line to a different system vibrating in a given special mode”, and that “if we were to take \( s = 2 \) in Eq. (25) and to attribute \( m \) integral values 3, 4, 5, ..., we should have a series of frequencies of the same general character as the hydrogen series, but still differing considerably in actuarial values”.

Now, Rayleigh points out a fundamental point as follows — “There is one circumstance which suggests doubts whether the analogue of radiating bodies to be sought at all in ordinary mechanical or acoustical systems vibrating
about equilibrium. For the latter, even when gyratory terms are admitted, gives rise to equations involving the square of the frequency; and its is only in certain exceptional cases, e.g. Eq. (25), that the frequency itself can be simply expressed.11 On the other hand, the formulae and laws derived from observation of the spectrum appear to introduce more naturally the first power of the frequency. For example, this is the case with Balmer's formula. Again, when the spectrum of a body shows several doublets, the intervals between the components correspond closely to a constant difference of frequency, and could not be simply expressed in terms of squares of frequency. Further, the remarkable law, discovered independently by Rydberg and Schuster, connecting the convergence frequencies of different series belonging to the same substance, points in the same direction."

At the end, Rayleigh points out the following — "What particular conclusion follows from this consideration, even if force be allowed to it, may be difficult to say. The occurrence of the first power of the frequency seems suggestive rather of kinetic relations than of those of dynamics." This should be said to be a really surprising excellent view. The nature here described was the reason why the spectral laws could not be treated by means of ordinary mechanics, and was the point that let quantum mechanics inevitably be arrived at. Ritz12 took in 1908 this problem pointed out by Rayleigh, to treat the combination rule. We shall describe it in Vol. II.

2. Ether Model of the Atom

§1. Atomism

It is well-known that the modern concept of atoms was introduced by Dalton at the beginning of the 19th century, mainly in connection with chemistry. Atoms were supposed to be solid bodies with impenetrability and rigidity, as is understandable from the formation of that concept, and also from the tradition at that time in philosophical thought as well as in mechanics. The atoms of various kinds of elements were assumed to have their respective characteristic masses. Though there was the idea that atoms were elastic bodies, or deformable bodies, it was not generally accepted. However, the supposition

11) Translator's note: Rayleigh's original number of Eq. (25) is (31).
of the inelasticity of atoms was still not unthinkable, because of the fact that atoms lost their kinetic energies in atomic collisions.

In order to overcome this difficulty R. G. Voscovich (1711–1787) thought for the first time such an image as is considered nowadays of atoms. In other words, he thought that atoms were no more than centers of attractive and repulsive forces acting at a distance, and were not such things as came into collision with each other at contact like ordinary solid bodies. A. M. Ampère, A. L. Cauchy and M. Faraday thought also that atoms were not of definite extension, but were mere centers of force.

It became one of the themes in the 19th century to think atoms to be mechanical solid bodies or mere centers of force. It became later a subject of the controversy between materialism mainly with mechanistic color and empiricismic idealism, and was influenced by their respective standpoints. Anyway, those who accepted atoms were on the whole materialistic. Empiricismic idealism did not further wish to accept atoms themselves, or the reality and importance of the concept of atoms. Thus, from the end of the 19th century to the beginning of the 20th, there arose acute disputes between atomism and energetics.

Kelvin estimated in 1883 the sizes of atoms to be $10^{-7}$ cm $\sim 10^{-8}$ cm. Boltzmann advocated passionately atomism. On the contrary, W. Ostwald denied still atomism even in 1904 when enough knowledge of atoms had been already accumulated. He claimed to make science free from hypothetical concepts of which no direct experimental verification could not be made, and called atomic and molecular theories harmful hypotheses, stressing direct study of experimental results. E. Mach denied also earnestly atomic and molecular theories. About this situation Cajori\textsuperscript{13}) mentions very appropriately that it is strange, that oppositions to atomism appeared just when primary facts which surely and decisively established the justice of atomism. Nowadays there is not anyone who denies still atomism, but epistemological arguments like Ostwald’s or Mach’s appear repeatedly. For one who goes with such a kind of idealism, the gap between his argument and his act in real study becomes more and more widely separated.

Now, let us return to the first problem. In order to overcome the difficulty of inelastic atoms, the supposition was made that elementary units of matter were not atoms, but were certain combinations of them, namely, molecules.

In the 19th century, however, the concept of molecules was used in various ways. Atoms were very often called molecules.

Molecules were treated as elastic bodies, even when atoms were supposed to be inelastic bodies. Clausius and Maxwell considered so, too. Kelvin\textsuperscript{14} connected the elasticity of molecules with the law of energy conservation, and said that we were denied the idea of the lack of full elasticity, namely, the inelasticity of ultimate molecules, by modern physical theory of the conservation of energy. On the basis of the idea that molecules would behave as perfectly elastic spheres in their collisions, the kinetic theory of gases was developed with a great success, by Joule (1848), Krönig (1856), Clausius (1857), Maxwell (1860), Boltzmann (1885), and others.

\section{Vortex atoms}

Another method of overcoming the difficulty of inelastic atoms was Kelvin's idea of vortex atoms.\textsuperscript{15} This was based on the standpoint of ether monism, and was the one that went further in the direction of clarifying electromagnetic phenomena by means of ether, which was supposed to have properties corresponding to those taken in the dynamics of elastic bodies or fluid. This way of thinking should be said to correspond to the present-day field monism, and aimed at making clear the fundamental nature of matter.

Kelvin investigated the similarity of atomic properties to fluid dynamical properties of vortex rings. Ether was supposed to be an incompressible fluid. It is not possible to give rigidity to an incompressible fluid at rest, nor to expect vibrations in it. The rigidity of atoms and their characteristic vibrations could not therefore be expected in this way. On the other hand, vortex rings such as rings of smoke coming out of the mouth of a smoker are elastic and have rigidity. Circles are their stable forms, deviations from which give rise to vibrations. When two vortex rings collide with or come close to each other, they make each other's course and form to bend and vibrate respectively.

Kelvin thought atoms of matter to be vortexes of ether. Atoms were supposed not to be foreign things in spreading ether, but to be parts or states of ether distinguished by their vortex motion. As vortexes, atoms should therefore have rigidity, move freely, have elasticity, collide with each other and make

\textsuperscript{14}Lord Kelvin, Phil. Mag. (4) 45 (1873), 329.
vibrations. Vortexes possessed all the properties of atoms except one. That was their gravity. It was not considered strange at that time, however, that the gravity of atoms could not immediately be treated, because this was very small.

According to Lodge,\textsuperscript{16} Thomson's theory of matter was however an unverified theory,\textsuperscript{17} and it took perhaps only a few step beyond a conjecture. But Lodge considered that it was worth knowing enough, and also worth using in work and study, though it might succeed or fail. He supposed that this idea was the simplest one of those which man thought of the universe of matter. According to him, there was in the universe the sole universal substance, which was completely homogeneous, had a continuous and simple structure, spread over as far as man knew, and existed equally everywhere. Some parts of it were at rest, or made simple irrotational motion, and transmitted the wave called light. The other parts made rotational motion, and if ether was a perfect fluid, vortexes in it did not diminish their energies and were perfectly separated from the rest on account of their motion. These parts constituted the thing which was called matter. They were thought to have rigidity.

Lodge said that a continuous substance filled up all the space, which vibrated as light, became positive and negative electricity on account of shear, and formed matter by making vortexes, and that it conveyed all the action and reaction worked by matter by means of its continuity instead of collision, these being ether and its functions.

Such an ether monism was not able, however, to treat the dynamical properties of ether consistently. Kelvin attempted to follow the properties of ether by means of various dynamical models, but ended in failure because of contradictions piling one upon another. He was not satisfied with the procedure of putting first some particular differential equations as in the Maxwell theory of electromagnetism. He wanted to reduce every thing mechanically to the Newton dynamics. In this regard he was a mechanical materialist. Such a way of thinking is similar to Einstein's, who cannot by any means accept quantum mechanics still now.

Kelvin's dynamical models were not such as to raise stepwise the question what substance existed in the object. In other words, they did not raise first the question what substance existed in atomic structure, but proposed the problem of developing dynamical models with properties corresponding to the

\textsuperscript{17}Thomson means here W. Thomson, that is, Lord Kelvin.
phenomena shown by atoms. The question whether the substantial structure of a dynamical model itself would be actually proved or not, was not taken to be an important one. This makes an essential difference between the method of dynamical models and the correct substantialistic method. The problem of atomic properties has been solved for the first time, after the question of what atoms consisted was raised and clarified.

3. Clarification by the Theory of Electrons

§1. Lorentz's theory of electrons

H. A. Lorentz had already in 1880 a clear thought of molecular structure, with which he derived the famous formula for the propagation of light in matter.\(^{18}\) In this paper, he says that Maxwell's theory was not yet applied to the propagation, reflection and refraction of light, and that he has derived here the relation between the refractive index \(n\) and the density of matter \(d\). Lorentz starts from the following idea, namely, the supposition of the existence of ether in the gaps among the molecules of matter, and the equality of the property of this medium to that of vacuum except the neighborhoods directly in contact with the molecules. He assumes also the following: In each molecule of isotropic matter an electric moment is induced by, in proportion to and in the direction of, the electromotive force, which obeys the Coulomb law at very close distances. By virtue of these assumptions, namely by the supposition of substantial molecules of matter separated from ether, it becomes possible to derive the relation between the specific inductivity \(\varepsilon\) of an isolator and its density \(d\), through calculations in which the electric polarization of the molecules of matter is taken into account. From this relation, together with the formula derived by Maxwell

\[
\varepsilon = n^2, \tag{27}
\]

Lorentz arrives at the conclusion that there must hold the relation

\[
\frac{n^2 - 1}{(n^2 + 2)d} = \text{const}. \tag{28}
\]

\(^{18}\)H. A. Lorentz, Wied. Ann. 9 (1880), 641. Cf. also ibid. 11 (1880), 70.
Thus, Lorentz's way of thinking is stepwise, and aims to give soundly the substantiality to atoms and molecules, instead of trying to clarify at once the whole problem of atoms and molecules.

Lorentz developed his idea more clearly in 1892 and on, to suppose that the fundamental charged particle had a mass, and applied Maxwell's theory of electromagnetism to it, to explain the optical and other electromagnetic phenomena of matter with a great success.\(^{19,20}\) He named the particle the ion, which means the electron nowadays. Every atom of matter was assumed to have ions, from the polarization and vibration of which the optical and electromagnetic properties were explained. The former idea of attributing these properties to the density of ether in matter was thus excluded, so that ether became simply the light ether, that is, the medium of light in vacuum.

In his work cited above,\(^{21}\) Lorentz says in the following way: In order to obtain the fundamental equations of the electric phenomena in moving bodies, he has participated in a hypothesis which is advocated by many physicists in recent years. In other words, he has assumed that in all the bodies there exist charged particles of small mass, and that all the electric processes depend on the positions and motions of these "ions". He mentions that in electrolysis this idea is considered the sole possible one, and that Giese,\(^{22}\) Schuster,\(^{23}\) Arrhenius,\(^{24}\) and Elester-Geitel\(^{25}\) observe the idea that the electric conduction in gases is the convection of ions. He then says that there seems no reason for preventing the idea, that the molecules of ponderable and dielectric matter have those particles, which are displaced from their positions only when an external electric force acts on them, realizing thereby the dielectric polarization of such matter.

The periodically vibrating polarization which produces light in Maxwell's theory becomes now the vibration of ions. In the old theory of light the dispersion of color was attributed to the resonance of ponderable matter. Lorentz considers that this holds good in outline also in the electromagnetic theory.
of light, but it is necessary to endow the ions with a mass. He says that he has shown this in a previous paper,\textsuperscript{26} but in that paper he has derived the equations of motion still from the action at a distance, and not from Maxwell’s concepts which he now considers much simpler, and that the same viewpoint has later been taken in Helmholtz’s electromagnetic theory\textsuperscript{27} of dispersion of color.\textsuperscript{28}

About the constitution of Lorentz’s theory of electrons, Lorentz says that the theory of ions has fitted his purpose very well, because it has enabled him to a satisfactory degree to introduce transparency into the equations for ether. The problem is thus divided into two. The first is the question how the state of ether, that is, the electromagnetic field in the present-day wording, is determined by the charge, position and motion of ions. The second is the question what force ether exerts on charged particles. As to the form of theoretical development, in his first work\textsuperscript{19} the derivation was carried out on a few assumptions by means of the d’Alembert principle, and a method was adopted similar to that with which Maxwell applied the Lagrange equations. In his later book\textsuperscript{20} Lorentz proceeded by taking the fundamental equations as the hypothesis.

The equations for ether (the equations of electromagnetic field) were supposed to agree with those in Maxwell’s theory. The change caused by an ion to ether was expressed as propagating with the light velocity. The force exerted by ether on a charged particle was thought to depend on the state of the medium at the position where the particle existed. This force set up for the first time by Lorentz is the one called the Lorentz force, and was successful in explaining the Zeeman effect. In the latter work under consideration, Lorentz mentions about the retardation of the propagation of force that the fundamental laws assumed are different in essential points from the laws set up by Weber and Clausius, so that the influence which particle B receives in the neighborhood of particle A is dependent on the motion of A, but not on the motion at that instant, being determined rather by the motion which A has made a time before. Lorentz says that the laws assumed correspond namely

\textsuperscript{26}H. A. Lorentz, Verhandlingen der Akad. van Wet. te Amsterdam, Deel 18 (1878); Wied. Ann. 9 (1880), 641.
\textsuperscript{27}H. Helmholtz, Wied. Ann. 48 (1893), 389.
\textsuperscript{28}Lorentz cites further the following papers: Koláček, Wied. Ann. 32 (1887), 224, 429; Goldhammer, Wied. Ann. 47 (1892), 93.
to the requirement which Gauss put on electrodynamics in his famous letter\textsuperscript{29} sent in 1845 to Weber.

Lorentz considers that his idea brings back in some sense the theory of electromagnetism in matter in the direction of the theory of electricity before Maxwell. Lorentz says that, though the core of Maxwell's idea is not lost in this process, it cannot be denied that the assumption of ions is near the idea of charged particles formerly treated, as is seen particularly clearly in simple cases. This means that, because every charge is thought to be generated by an aggregation of positive and negative charges, and because the fundamental equation of ions is to give the Coulomb law, electrostatics must become identical in form with the theory before Maxwell. Thus, Lorentz adopts the formulation of the interaction of the electromagnetic fields obeying the Maxwell equations with charged particles, abandoning the electromagnetics based on the density of ether and the phenomenological theory by Maxwell of ether in matter.

The introduction of the force which an electromagnetic field exerts on a charged particle, that is, the Lorentz force is due to Lorentz. Let us describe how it is done, (using the electromagnetic system of units).

Maxwell considered two kinds of deviations from a state of equilibrium in the medium of ether. The first is the one which occurs in the neighborhood of a charged particle, that is, the electric displacement $d$. [Lorentz calls it the dielectric movement (\textit{dielektrische Verschiebung}).] It distributes solenoidally in the space of pure ether between ions.\textsuperscript{30}

\begin{equation}
\text{div } d = 0 .
\end{equation}

Let the space density of electric charge be denoted by $\rho$. Equation (29) does not hold in the region where $\rho$ is different from zero. We have there

\begin{equation}
\text{div } d = \rho ,
\end{equation}

from Maxwell's theory.

The second kind of deviation is determined by the magnetic force $H$, which satisfies the following equations in relation to the density of current $J$ at the instant under consideration.

\begin{equation}
\text{div } H = 0 ,
\end{equation}

\textsuperscript{29}K. F. Gauss, Werke, Bd. 5, p. 629.
\textsuperscript{30}Ref. 20), p. 14.
Clarification by the Theory of Electrons

\[ \text{rot } \mathbf{H} = 4\pi \mathbf{J}. \]  
(32)

In terms of the light velocity \( c \) the Maxwell equation in vacuum is\(^{31}\)

\[-4\pi c^2 \text{rot } \mathbf{d} = \dot{\mathbf{H}}.\]  
(33)

If ponderable matter is at rest, so that \( \mathbf{d} \) is independent of the time, we have \( \mathbf{J} = \mathbf{H} = 0 \). \( \mathbf{d} \) satisfies then the equations

\[ \text{div } \mathbf{d} = \rho, \]  
(34)

\[ \text{rot } \mathbf{d} = 0. \]  
(35)

The components of \( \mathbf{d} \) is therefore the respective partial derivatives of a single function. If this function is denoted by \(-\omega/4\pi\), we have

\[ d_x = -\frac{1}{4\pi} \frac{\partial \omega}{\partial x}, \ldots. \ldots \]  
(36)

Inserting these into Eq. (34), we get the Poisson equation

\[ \Delta \omega = -4\pi \rho. \]  
(37)

The first force exerted on ponderable matter\(^ {32}\) — The former electrostatics was in accordance with experience. According to that, the force exerted on a volume element \( d\tau \) is derived by obtaining first the potential function from the Poisson equation, and multiplying then its gradient with \(-c^2 \rho d\tau\). Because the equation given for \( \omega \) is the Poisson equation, the potential function should be identical with \( \omega \). The components of the force are accordingly given by

\[ -c^2 \rho \frac{\partial \omega}{\partial x} d\tau, \ldots. \ldots \]  
(38)

As in Maxwell's theory, the force should be attributed to the state of ether. It depends therefore on the electric displacement. From the relation between \( \omega \) and \( \mathbf{d} \) we have for the components of the force

\[ 4\pi c^2 d_x \rho d\tau, \ldots. \ldots \]  
(39)

---

\(^{31}\) \( \dot{\mathbf{H}} = d\mathbf{H}/dt \), being the time.

\(^{32}\) Ref. 20), p. 18.
This shows that, as long as the electric displacement occurs in a volume element $d\tau$, ether exerts the force given above on the ponderable matter which exists there. About a unit charge the force is expressed as

$$F_1 = 4\pi c^2 d.$$  \hspace{1cm} (40)

Let there be two ions of charge $e$ and $e'$ respectively. Their sizes are assumed to be small compared with the distance between them. For the purpose of obtaining the force exerted on the first ion, its volume is divided into elements, and the use is made of Eq. (39) to be integrated. $d$ is the sum of those due to both the ions, but $d$ due to the first ion is assumed to exert no force on the first particle itself. The force due to the second ion is, everywhere inside the first ion, in the direction of the distance $r$ between both the ions, and is of the strength $e'/4\pi r^2$, so that $e$ is exerted by $e'$ the force

$$\frac{e^2 ee'}{r^2}.$$  \hspace{1cm} (41)

This is in agreement with the Coulomb law. Lorentz concludes thus that the ion theory is in agreement with the electrostatics before Maxwell.

The second force exerted on ponderable matter$^{33)}$ — If $u$ is the velocity of each ion and $I$ is the volume in a tube of flow surrounded by two sides of the area $\theta$ being at the relative distance $ds$, we have

$$I = \theta ds,$$  \hspace{1cm} (42)

and for the average density of current $\overline{J}$ when the charge moves with the velocity $u$

$$\overline{J} = \frac{1}{I} \int \rho u_x d\tau, \ldots, \ldots.$$  \hspace{1cm} (43)

In the case of ions this becomes

$$\overline{J} = \frac{1}{I} \sum \rho u_x, \ldots, \ldots.$$  \hspace{1cm} (44)

Because the current $i$ is given by

$$i = \theta \overline{J},$$  \hspace{1cm} (45)

$^{33)}$Ref. 20), p. 20.
we have
\[ \sum e\mathbf{u} = ids . \] (46)

*ids* is a vector in the direction of the current.

Let us consider an element of current in an external magnetic field. As is well known the force exerted on it is
\[ [ids, \mathbf{H}] . \] (47)

Because of Eq. (46) this is equal to
\[ \left[ \sum e\mathbf{u}, \mathbf{H} \right] = \sum \{ e[\mathbf{u}, \mathbf{H}] \} . \] (48)

Since this consists of the forces exerted on each ion, the force exerted on an ion can be considered to be
\[ e[\mathbf{u}, \mathbf{H}] . \] (49)

Generally speaking, in the case of a continuous distribution of charge, it is given by
\[ \rho d\tau [\mathbf{u}, \mathbf{H}] . \] (50)

For a unit charge we have
\[ F_2 = [\mathbf{u}, \mathbf{H}] . \] (51)

We see thus that a unit charge is under the same electrodynamical action as that which is exerted on a conductor transmitting an electric current.

Through the composition of \( F_2 \) with \( F_1 \) given above, the total force exerted on a unit charge, which is called the electric force, becomes
\[ \mathbf{F} = 4\pi e^2 d + [\mathbf{u}, \mathbf{H}] . \] (52)

By this force the motion of ions is determined.\( ^{34} \)

Let us then obtain the field generated by moving ions.\( ^{35} \) From Eq. (32) we have
\[ \text{rot rot} \mathbf{H} = \text{rot}(4\pi \mathbf{J}) . \] (53)

\( ^{34} \)Lorentz notes in this connection that every ion is assumed not to make rotation.

\( ^{35} \)Ref. 20), p. 50.
Because

\[ 4\pi J = 4\pi \rho u + 4\pi d, \]  

(54)

we get

\[ c^2 \Delta H_x - \left( \frac{\partial^2 H_x}{\partial t^2} \right) = 4\pi c^2 \left[ \frac{\partial (\rho u_y)}{\partial z} - \frac{\partial (\rho u_z)}{\partial y} \right], \ldots, \ldots \]  

(55)

If the three functions \( \varphi_x, \varphi_y, \varphi_z \) satisfy respectively the equations

\[ c^2 \Delta \varphi_x - \frac{\partial^2 \varphi_x}{\partial t^2} = 4\pi c^2 \rho u_x, \ldots, \ldots, \]  

(56)

we obtain

\[ H_x = \left( \frac{\partial \varphi_y}{\partial z} \right) - \left( \frac{\partial \varphi_z}{\partial y} \right), \ldots, \ldots. \]  

(57)

If we put

\[ \rho u_x = f(x, y, z, t), \]  

(58)

we have as the solution of the above equation for \( \varphi_x \)

\[ \varphi_x = -\int \frac{1}{r} f \left( \xi, \eta, \zeta, t - \frac{r}{c} \right) d\tau, \]  

(59)

where the point at which the strength of the field is to be obtained is \( P(x, y, z) \), and the integration is to be carried out with respect to the point \( Q(\xi, \eta, \zeta) \) at which \( \rho u_x \) does not vanish, \( r \) being \( \overline{PQ} \).

Consider now that a molecule, which has ions capable of oscillating, makes an electric oscillation. If \( Q_0 \) is a point in the molecule, \( P \) a point at a long distance compared with the size of the molecule, and \( r_0 = \overline{Q_0P} \), we have

\[ \varphi_x = -\frac{1}{r_0} \int \rho u_x d\tau, \]  

(60)

where the time at \( Q_0 \) is to be taken as \( t - r_0/c \). If \( u_x \) of every point in an ion is the same, the integral becomes \( \sum c u_x \). Denoting the displacement of an ion from its point of equilibrium by \( q \), we have

\[ u_x = \frac{dq_x}{dt}, \]  

(61)
and accordingly

$$\sum e\mu_x = \frac{d}{ct} \sum e\rho x.$$  \hspace{1cm} (62)

$\sum e\rho$ is called the electric moment of a molecule and is denoted by $m$, so that

$$\sum e\rho x = m_x,$$  \hspace{1cm} (63)

$$\varphi_x = -\frac{1}{r_0} \frac{dm_x}{dt} = -\frac{\partial}{\partial t} \left( \frac{m_x}{r_0} \right).$$  \hspace{1cm} (64)

Assuming that the oscillation in the molecule is represented by

$$m_x = a \cos 2\pi \frac{t}{T}, \quad m_y = 0, \quad m_z = 0,$$  \hspace{1cm} (65)

we have at the distance $r$

$$H_x = 0, \quad H_y = \frac{\partial}{\partial t} \left( \frac{\partial \chi}{\partial z} \right), \quad H_z = -\frac{\partial}{\partial t} \left( \frac{\partial \chi}{\partial y} \right),$$  \hspace{1cm} (66)

$$\chi = \frac{a}{r} \cos \frac{2\pi}{T} \left( t - \frac{r}{c} \right).$$  \hspace{1cm} (67)

Lorentz states that this is in accordance with that given by Hertz for the oscillation in the neighborhood of an oscillator.36)

After Lorentz’s first theory of electrons, Larmor criticized also the ether theory of matter from the standpoint of the theory of electrons, and made clear the necessity of introducing the definite substance of charged particle of matter, which he called the electron.37)

§2. The Zeeman effect

Mathematical descriptions such as Balmer’s formula were phenomenological ones, and could not provide sufficient knowledge for the mechanism of the radiation from atoms. It became possible, however, to take a step into the mechanism of radiation, when spectral lines were varied by changing the condition of radiating atoms. In 1897 the variation in spectral lines radiated from atoms in a magnetic field was discovered by Zeeman.38)


37) J. Larmor, Phil. Trans. A (1894), 812; A (1895), 713.

38) P. Zeeman, Phil. Mag. (5) 43 (1897), 226.
About this discovery there is an episode. Several years before the discovery, in the course of his measurements of the Kerr effect, it occurred to him whether the light of a flame if submitted to the action of magnetic field would perhaps undergo any change. With an extemporary apparatus the spectrum of a flame, colored with sodium, placed between the poles of an electromagnet, was looked at, but the result was negative. In the paper cited above, Zeeman says — “Probably I should not have tried this experiment again so soon had not my attention been drawn some two years ago to the following quotation from Maxwell’s sketch of Faraday’s life.” The quotation reads — “…we may mention that in 1862 he made the relation between magnetism and light the subject of his very last experimental work. He endeavoured, but in vain, to detect any change in the lines of the spectrum of a flame when the flame was acted on by a powerful magnet.” About this Zeeman thinks — “If a Faraday thought of the possibility of the above-mentioned relation, perhaps it might be yet worth while to try the experiment again with the excellent auxiliaries of spectroscopy of the present time, as I am not aware that it has been done by others.”

Zeeman used an electromagnet of medium size. The current furnished by accumulators was in most of cases 27 A, and could be raised to 35 A. The light was analysed by a Rowland grating, with a radius of 10 ft. and with 14,938 lines per inch. The first spectrum was used.

Between the paraboloidal poles of the magnet the middle part of the flame from a Bunsen burner was placed, the distance between the poles being 7 mm. “A piece of asbestos impregnated with common salt was put in the flame in such a manner that the two D-lines were seen as narrow and sharply defined lines on the dark ground. If the current was put on, the two D-lines were distinctly widened. If the current was cut off they returned to their original position. The appearing and disappearing of the widening was simultaneous with the putting on and off of the current. The experiment could be repeated an indefinite number of times.” With the red lines of lithium wholly analogous phenomena were observed. One may reason that, under the influence of the magnet, the temperature and also the density of the flame are increased. On this point, the absorption spectrum of the D-lines of natrium was examined, and the same result was obtained. Also in this case, the change in the density

39) It was in 1883 that H. A. Rowland in U.S.A. used a concave grating in his study of the spectrum of the sunlight.
due to convection currents caused by the magnetic field was suspected, and endeavors were made to eliminate this effect.

In this way Zeeman makes it clear that the spectral lines of incandescent vapour are widened by the action of magnetism, and says that "if this is really the case, then by the action of magnetism on the free vibrations of the atoms, which are the cause of the line-spectrum, other vibrations of changed period must be superposed".

Zeeman considers that, in the case of a band-spectrum this effect does not appear, since the mechanism of lighting is different not only qualitatively but also quantitatively. He says in accordance with Kayser's consideration — "In the case of a band-spectrum the molecules are complicated; in the case of a line-spectrum the widely separated molecules contain but a few atoms." He made experiments on the absorption spectrum of iodine, without any result recognized of the effect of magnetism. He concludes conversely from this that convection currents are insufficient to cause a change of the absorption spectrum of the D-lines.

Saying "Although the means at my disposal did not enable me to execute more than a preliminary approximate measurement, I thought it of importance to determine approximately the value of the magnetic change of the period", Zeeman estimates from the widening of the sodium lines which amounted to about 1/40 of the distance between the lines, and the intensity of the magnetic filed of about $10^4$ c.g.s., the positive and negative magnetic change to be of 1/40,000 of the period.

The reasoning by which Zeeman was induced to search after the influence of magnetism, was at first the following — If the hypothesis is true that in a magnetic field a rotatory motion of the ether is going on, the axis of rotation being in the direction of the magnetic forces (Kelvin and Maxwell), and if the radiation of light may be imagined as caused by the motion of the atoms, relative to the center of mass of the molecule, revolving in all kinds of orbits, suppose for simplicity circles; then the period, or, what comes to the same, the time of describing the circumstance of these circles, will be determined by the forces acting between the atoms, and then deviations of the period to both sides will occur through the influence of the perturbing forces between ether and atoms. The sign of the deviation of course will be determined by the direction of motion, as seen from along the line of force. The deviation will

40) The Zeeman effect of band-spectra was found in 1905 by R. W. Wood.
be the greater the nearer the plane of the circle approximates to a position perpendicular to the line of force.

Later Zeeman elucidated this subject by representing to himself the influence exercised on the period of a vibrating system if this is linked together with another in rapid rotatory motion. He compared this problem with that of a rotating cord treated by Kelvin in 1856. By supposing that the double pendulum is substituted by a luminiferous atom and the rotating arm by the rotational motion about the magnetic lines of force, Zeeman considers that the relation of the mechanical problem to the case of a luminiferous atom in a magnetic field will be clear.

§3. Explanation by Lorentz’s theory of electrons

Zeeman says that though the above-mentioned mechanical considerations were the first motive of his experiments, to him “a real explanation of the magnetic change of the period seemed to follow from Prof. Lorentz’s theory”. According to Lorentz’s theory Zeeman considers in the following way — In this theory it is assumed that in all bodies small electrically charged particles with a definite mass are present, that all electric phenomena are dependent upon the configuration and motion of these “ions”, and that light-vibrations are vibrations of these ions. Then the charge, configuration, and motion of the ions completely determine the state of the ether. The said ion, moving in a magnetic field, experiences mechanical forces of the kind above mentioned, and these must explain the variation of the period. Zeeman states — “Prof. Lorentz, to whom I communicated these considerations, at once kindly informed me of the manner in which, according to his theory, the motion of an ion in a magnetic field is to be calculated, and pointed out to me that, if the explanation following from his theory be true, the edges of the lines of the spectrum ought to be circularly polarized. The amount of widening might then be used to determine the ratio between charge and mass, to be attributed in this theory to a particle giving out the vibrations of light”. This extremely remarkable conclusion of Lorentz relating to the state of polarization in the magnetically widened lines was fully confirmed immediately by Zeeman’s experiments.

Now, Zeeman describes in his paper under consideration the manner of calculation informed by Lorentz. An ion is assumed to vibrate in the plane of \((x, y)\) in a uniform magnetic field, in which the magnetic force is everywhere parallel to the axis of \(z\) and equal to \(H\). Let \(e\) be the charge (in electromagnetic
measure) of the positively charged ion, and \( m \) its mass. The equations of motion then are

\[
\begin{align*}
    \frac{d^2 x}{dt^2} &= -k^2 x + eH \frac{dy}{dt}, \\
    \frac{d^2 y}{dt^2} &= -k^2 y - eH \frac{dx}{dt},
\end{align*}
\]

(68)

to which Zeeman adds a footnote, saying “These equations are like those of the Foucault pendulum, and of course lead to similar results”.

The first term of the second member expresses the elastic force drawing back the ion to its position of equilibrium, and the second term gives the mechanical force due to the magnetic field. These equations are satisfied by

\[
\begin{align*}
    x &= \alpha e^{\alpha t}, \\
    y &= \beta e^{\alpha t},
\end{align*}
\]

(69)

provided that

\[
\begin{align*}
    ms^2 \alpha &= -k^2 \alpha + eH s \beta, \\
    ms^2 \beta &= -k^2 \beta - eH s \alpha,
\end{align*}
\]

(70)

where \( m, k \) and \( e \) are to be regarded as known quantities.

The period \( T \) is particularly interesting. If \( H = 0 \), it follows from Eq. (70) that

\[
s = i \frac{k}{\sqrt{m}} = i \frac{2\pi}{T}
\]

(71)
or

\[
T = \frac{2\pi \sqrt{m}}{k}.
\]

(72)

In the case of \( H \neq 0 \), it follows approximately that

\[
s = i \frac{k}{\sqrt{m}} \left( 1 \pm \frac{eH}{2k\sqrt{m}} \right).
\]

(73)

Putting \( T' \) for the period in this case, we have

\[
T' = \frac{2\pi \sqrt{m}}{k} \left( 1 \pm \frac{eH}{2k\sqrt{m}} \right).
\]

(74)
Hence the ratio of the change of period to the original period becomes

\[ \frac{eH}{2k\sqrt{m}} = \frac{e}{m} \cdot \frac{HT}{4\pi} \]  \hspace{1cm} (75)

A particular solution of Eq. (68) is that representing the motion of the ions in circles. If revolving in the positive direction (viz., in the direction of the hands of a watch for an observer standing at the side towards which the lines of force are running) the period is somewhat less than if revolving in the negative direction. The period in the first case is determined by the value of Eq. (74) with the minus sign, in the second with the plus.

The general solution of Eq. (68) shows that the ions describe, besides circles, also slowly rotating elliptical orbits. In the general case, the original motion of the ion has an arbitrary position, and the projection of the motion in the plane of \((x, y)\) has the same character. The motion resolved in the direction of the axis of \(z\) is a simple harmonic motion, independent of and not disturbing the one in the plane of \((x, y)\), and hence is not influenced by the magnetic forces.

Imagine an observer looking at a flame placed in a magnetic field in such a direction that the lines of force run towards or from him. The observer will see the ions under consideration in the following way. There are some ions moving in circles and hence emitting circularly polarized light; if the motion is round in the positive direction the period will, for instance, be longer than with no magnetic field; if in the negative direction, shorter. There will also be ions seemingly stationary and really moving parallel to the lines of force with unaltered period. In the third place there are ions which seem to move in rotating elliptical orbits.

Zeeman makes use of a rule deduced from Lorentz's theory, according to which the state of the ether at a very great distance from a molecule will be obtained by superposing the states originated by every component of the motion of the moving ions. It follows in the first place that a circular motion of an ion gives circularly polarized light to points on the axis of the circle. Further, one may choose instead of the above considered elliptical orbits a resolution more suited to our purpose. One may resolve the motion of the ion, existing before the putting on of the magnetic force, into a rectilinear motion parallel to the axis of \(z\) and two circular motions in the plane of \((x, y)\). The first remains unchanged under the influence of the magnetic force, but the periods of the last are changed.
By the grating the vibrations originated by the motion of the ions are sorted according to the period, and the complete motion is broken up into three groups. The line will be a triplet. At any rate one may expect that the line of the spectrum will be wider than in the absence of the magnetic field, and that the edges will give out circularly polarized light. Zeeman notes that Stoney endeavours to explain the existence of doublets and triplets in a spectrum by the rotation of the elliptical orbits of the "electrons" under the influence of perturbing forces.

§4. Experimental verification by Zeeman of Lorentz's theory

Zeeman points out now that "a confirmation of the last conclusion may be certainly taken as a confirmation of the guiding idea of Prof. Lorentz's theory". To decide this point by experiment, an electromagnet with pierced poles was placed so that the axes of the holes were in the same straight line with the center of the grating. The quarter-undulation plate and nicol were placed between the grating and the eyepiece. The plate and the nicol were placed relatively in such a manner that right-handed circularly polarized light was quenched. According to the preceding the widened line must at one edge be right-handed circularly polarized, and at the other edge left-handed. By a rotation of the analyser over 90° the light that was first extinguished will be transmitted, and vice versa. Or, if first the right edge of the line is visible in the apparatus, a reversal of the direction of the current makes the left edge visible. Zeeman proceeded with the experiment in this way. "The cross-wire of the eyepiece was set in the bright line. At the reversal of the current the visible line moved! This experiment could be repeated any number of times". As a small variation of the preceding experiment, the analyser was turned round with the position of the quarter-wave plate unchanged. "The widened line is then, during one revolution, twice wide and twice fine".

The electromagnet was turned 90° in a horizontal plane, so that the lines of force became perpendicular to the line joining the slit with the grating. "The edges of the widened line now appeared to be plane polarized, at least in so far as the present apparatus permitted to see, the plane of polarization being perpendicular to the line of the spectrum. This phenomenon is at once evident from the above consideration. The circular orbits of the ions being perpendicular to the lines of force are now seen on their edges".
The use of such polarization proves that the line is broken into three lines, though the line is not apparently seen to be so broken, and further confirms the expectation from Lorentz's theory of "the light vibrations caused by the motion of ions".

In addition, there is obtained an important quantitative result. That is, from the measured widening of the period of the line of $1/46,000$, the value of $e/m$ is deduced according to Eq. (47). Zeeman states — "It thus appears that $e/m$ is of the order of magnitude $10^7$ electromagnetic c.g.s. units." This is very well compared with the present-day value $e/m = 1.7 \times 10^7$, and should be said to give the first real or substantial foundation of the theory of electrons.\footnote{Shortly after Zeeman, J. J. Thomson measured $e/m$ using cathode rays for the first time, and gave various values between $(0.6-2) \times 10^7$ to it [Phil. Mag. (5) 44 (1897), 293].}

Zeeman also says — "It may be deduced from the above experiment whether the positive or negative ion revolves. If the lines of force were running towards the grating, the right-handed circularly polarized rays appeared to have the smaller period. Hence in connection with the above equations it follows that the positive ions revolve,\footnote{This is of course the mistake of the negative ions. Zeeman corrected it in a footnote to his paper published next [Phil. Mag. (5) 44 (1897), 55].} or at least describe the greater orbit."

Zeeman concludes his paper by saying that "it seems very promising to investigate the motion of the ions for various substances, under varying circumstances of temperature and pressure, with varying intensity of magnetization", and by mentioning the possibility of knowing the intensity of the magnetic force at the surface of the sun from the spectrum of sunlight.

This experiment was noticed by various persons and experiments were carried out at various places. Oliver Lodge\footnote{O. Lodge, The Electrician, Feb. 28, 1897, p. 567, as cited by Zeeman in Ref. 44).} doubted the interpretation of this experiment. Lodge made an experiment himself and saw the doubling of the lines, but did not suppose it to be a real doubling of the lines, thinking it to be a broadening and a reversal. In regard to these questions, Zeeman\footnote{P. Zeeman, Phil. Mag. 44 (1897), 55.} made careful experiments to verify Lorentz's theory.

In this work, he did not succeed completely about sodium in observing with the means at his disposal the expected doublets and triplets. He succeeded at last in observing with the blue line ($\lambda = 480 \mu\mu$) of cadmium the doublet and the triplet. It had already been remarked by Egoroff and Georgiewsky\footnote{Egoroff and Geogiewsky, Compt. Rend. t. cxxiv. pp. 648, as cited by Zeeman in Ref. 44).} that
cadmium also exhibited the broadening of the spectral lines under the action of magnetism. Zeeman used a Rowland grating with a radius of 6 ft. It had 14,438 lines to the inch, giving a satisfactory second spectrum. A spark was used between cadmium electrodes, and the cadmium line was especially sharp at the violet side.

Zeeman's experiment verified completely Lorentz's theory. Zeeman says in his paper under consideration about Lodge's opposition to this result—"Lodge\textsuperscript{46}) has even pointed out in an extremely interesting paper that my value of $e/m$ can be explained without the necessary motion of any 'matter' at all."

Lodge informed Zeeman afterward that he had seen the doubling of the lines, but he did not suppose it to be a real doubling of the lines, thinking it to be due to a broadening and an absorptive reversal of the central part. Against Lodge's opinion, Zeeman mentioned the observation by himself of the perpendicular polarizations of the middle and of the edges of the triplet. It should be very noticeable from the viewpoint of methodology that Zeeman pointed out the fact, that Lodge had not made the observation of this last phenomenon, and Lodge's object was not in the first place to test Lorentz's theory.

§5. Measurement of $e/m$

A more exact measurement of $e/m$ by means of the Zeeman effect was then carried out by Zeeman.\textsuperscript{47)} With the use of the magnetic field of 32,000 c.g.s. and without a nicol, the cadmium line was broken up into three lines, separated by dark spaces, so that the magnitude of the change due to the magnetic field could be exactly measured. If the central line was quenched by means of a nicol and the distance between the side lines was measured, it corresponded to the double change of the period. This distance was measured with the use of a micrometer eyepiece to increase greatly the accuracy of measurement. This measurement was done at the laboratory of Prof. Haga of the University of Groningen. The grating used had a radius of 10 ft. and 10,000 lines to the inch. The source from which the light of sodium was emitted, was a piece of asbestos paper soaked with molten salt and introduced into the flame of coal-gas fed with oxygen under high pressure. With the placement of nicol and the putting on of the current, the distance between the two lines, namely, the position of

\textsuperscript{46}) O. Lodge, The Electrician, Mar. 12, 1897, as cited by Zeeman in Ref. 44).

\textsuperscript{47)} P. Zeeman, Phil. Mag. \textbf{44} (1897), 255.
each of $2 \times 2$ lines formed by the D-lines was read. The magnetic change was found to be the same for both the D-lines. The intensity of the magnetic field was 22,000 c.g.s., and the positive and negative magnetic change of the period amounted to $1/17,800$. Hence it resulted that $^{48)}$ 

$$\frac{e}{m} = 1.6 \times 10^7.$$ 

This value was in close agreement with the result of his previous rough measurement.

§6. **Larmor’s theory**

After Zeeman’s study Michelson$^{49)}$ and Lodge$^{50)}$ made experiments and confirmed the Zeeman effect.

Larmor$^{51)}$ generalized Lorentz’s theory of this effect to throw light on its nature by means of the method of coordinate transformation, and showed that this effect could take place for the case of general kinds of force other than the force proportional to the distance from the position of equilibrium.

In this paper, Larmor considers first the simple case of a single ion of charge $e$ and effective mass $M$, describing an elliptic orbit under an attraction to a fixed center proportional to the distance therefrom. The equations of motion are$^{52)}$

$$\ddot{x}, \ddot{y}, \ddot{z} = -a^2(x, y, z), \quad (76)$$

where $a/2\pi$ is the frequency of oscillation in any direction. Now suppose that a uniform magnetic field $H$, in a direction $(l, m, n)$, is introduced. The equations of motion (in electrostatic units) become

$$\begin{align*}
\dot{x} &= -a^2 x + \kappa(ny - mz), \\
\dot{y} &= -a^2 y + \kappa(lz - nx), \\
\dot{z} &= -a^2 z + \kappa(mx - ly),
\end{align*} \quad (77)$$

---

$^{48)}$ In Zeeman’s paper one reads $1.6 \times 10^{-10}$ for the ratio. This is of course a mistake.

$^{49)}$ A. A. Michelson, Phil. Mag. May, 1897.


$^{51)}$ J. Larmor, Phil. Mag. **44** (1897), 503.

$^{52)}$ The dot means $d/dt$, so that $\ddot{x} = d^2x/dt^2$. 
where $\kappa = eH/Mc^2$, $c$ being the light velocity. To obtain the frequencies $p/2\pi$ thus modified by the magnetic field, $(x, y, z)$ are put to be proportional to $e^{ipt}$. This gives the equation for $p$

$$(p^2 - a^2)^3 - \kappa^2 p^2 (p^2 - a^2) = 0.$$ 

Corresponding to each original period represented by $p = a$, there are three modified ones represented by

$$\begin{cases}
p = a, \\
p^2 \pm \kappa p - a^2 = 0.
\end{cases}$$

When, as in practice, $\kappa$ is very small, the two latter will be approximately $p = a + \kappa^2/8a \pm \kappa/2$, or with $\kappa^2$ neglected further

$$p = a \pm \kappa/2.$$ (78)

One thus obtains triple lines. Larmor says that "the striking feature is that the modification thus produced is the same whatever be the orientation of the orbit with respect to the magnetic field".

Larmor makes then the generalization of this result. Suppose that the original orbit is referred to a system of axes $(x, y, z)$, that are themselves revolving with angular velocity $\omega$ round an axis of which the direction is $(l, m, n)$. The component velocities $(u, v, w)$ referred to this moving space are

$$\dot{x} - y\omega n + z\omega m, \ldots,$$

and the component accelerations are

$$\ddot{u} - v\omega n + w\omega m, \ldots.$$ 

Thus the component acceleration parallel to $x$ is

$$\ddot{x} - 2\omega (ny - mz) - \omega^2 x + \omega^2 l(lx + my + nz).$$

If $\omega$ is so small that $\omega^2$ can be neglected, this becomes

$$\ddot{x} - 2\omega (ny - mz).$$

On the other hand, the same quantity in the previous simple case is from Eq. (77)

$$\ddot{x} - \kappa (ny - mz) = -a^2 x.$$
If, then, we take
\[ \omega = \frac{\kappa}{2}, \]
we have the acceleration identical with that in the simple case. "In other words, the oscillation thus modified will be brought back to its original aspect if the observer is attached to a frame which revolves with angular velocity \( \kappa/2 \) or \( eH/2Mc^2 \) round the axis of the magnetic field. In a circular orbit described one way round this axis the apparent rotation will in fact be retarded, in one described the other way round it will be accelerated, in a linear oscillation along the axis there will be no alteration: hence the three periods found above", as Larmor points out. \( eH/2mc^2 \) is called nowadays the Larmor frequency, and the rotational motion caused by the magnetic field the Larmor precession.

Larmor points out that the argument given above still applies, whatever be the number of revolving ions in the molecule, and however they attract each other or are attracted to fixed centers on the axis, provided \( \kappa \) has the same value for them all, in any such case the actual oscillation in the magnetic field being identical with the unmodified oscillation as seen from a revolving frame.

He remarks also that the study of various spectral lines makes it possible to investigate those which move in atoms or molecules. According to him, a view has been enunciated that they are only negative ions that are free to vibrate in the atom or molecule. And, on such a hypothesis, if the charges of these negative ions are proportional to their effective masses, for example, if they are simple electrons without inertia than that of the electric charge, the intervals (measured in difference of frequency) between magnetic doublets and the outside lines of magnetic triplets in the spectrum should be the same for all lines. Moreover, they should be the same in different spectra. Larmor states that "thus a hypothesis of that kind can be definitely put to the test".

He continues to discuss that when there are ions of different kinds describing orbits in the molecule, these exact results no longer hold, but even then we can assert that the difference of frequency between the lines of a magnetic doublet is of the order of \( eH/2\pi Mc^2 \), and the order of magnitude of \( e/M \) can be thence derived. Larmor states — "Thus Zeeman concludes from his experiments that the effective mass of a revolving ion, supposed to have the full unitary charge or electron, is about \( 10^{-3} \) of the mass of the atom. This is the same as Professor J. J. Thomson's estimates of the masses of electric
carriers in the cathode rays. If we took these carriers to be simply electrons, as their constancy under various environments tends to indicate, there would thus be about $10^3$ electrons in the molecule.” Larmor’s statement should be noted as the one that shows the idea at that time of the atom or molecule, together with Thomson’s.

Larmor treats then the vibration of a system of electrons or ions, confined to a surface over which they are free to move and constituting an electric charge on it, and shows that the oscillations of the form of a surface of revolution are modified by the introduction of a magnetic field, just as if the angular velocity $\kappa/2$ were imparted to the vibrating system. He applies this consideration to the explanation of the series of double lines in a spectrum. His explanation goes as follows — “Suppose, as a very rough illustration, that a polar molecule is constituted of a system of positive electrons around one pole and a system of negative ones around the other, the two systems being so far apart as to have practically separate sets of periods for their orbital motions, each of course disturbed by the presence of the other. Each of these systems moves in the magnetic field, more or less constant, arising from the other; and the effect of this disturbing field will, as above, be to duplicate all the periods of that set in the above regular way.”

Larmor formulates the relation between the motion of the electrons and the radiation emitted by them, according to the Maxwell–Lorentz theory of electromagnetism. He treats the modes of electric vibration, such as the polarization of light in the Zeeman effect, for example. The specification of that radiation is assigned by a summation over the different elements of the paths of the oscillating ions. Suppose that an ion $e$ is at $A$ and after a time $\delta t$ at $B$, where $\overline{AB} = v\delta t$, $v$ being its velocity. The effect of this displacement is the same as that of the creation of an electric doublet $\overline{AB}$ of moment $M = ev\delta t$. We have to find the influence propagated from such a doublet, and then integrate along the paths of the ions of the molecule.

Consider now such a doublet at the origin, lying along the axis of $z$. The lines of magnetic force will be circles around the axis, and the force will be specified by a single variable, its intensity $H$. Let $r$ be the distance from the origin to a point at which $H$ is to be calculated, and $\theta$ the angle between the line connecting the origin with the point and the axis of $z$. The doublet is

---

53) J. J. Thomson’s work cited here was published in Phil. Mag. 44 (1897), 293, shortly before Larmor’s paper under consideration.
treated as a linear current-element of strength \( dM/dt \). At a point in the \((x-y)\) plane close to such an element, \( H \) is given by \( H = -r^{-2}dM/dt \).

On the other hand, the solution of the wave equation is given in the form

\[
H = -\sin \theta \left( \frac{f(t - r/c)}{r^2} + \frac{f'(t - r/c)}{cr} \right).
\]

Because this becomes \( H = -r^2 f(t) \) when \( \theta = \pi/2 \) and \( r \) is very small, it follows that

\[
\frac{dM}{dt} = f(t).
\]

The values of \( H \) at any arbitrary point is therefore obtained from Eq. (79), if \( f(t) \) is so given as to represent the moment of the doublet. Larmor considers then the effect of suddenly establishing the doublet \( M = ev\delta t \) at the origin. There will be a thin spherical shell of magnetic force propagated out with the velocity \( c \), the total force integrated across the shell being exactly \(-Mr^2 \sin \theta\), for the integral of the second term in \( H \) vanishes. This amount of magnetic force is the same as the magnetic force due to a permanent steady current-element of intensity \( M/\delta t \), or \( ev \).

Larmor arrives thus at the following conclusion. The magnetic force at a point at distance \( r \) due to a moving ion depends on the state of the ion at a time \( r/c \) previously; for near points it is in the plane perpendicular to \( r \), at right angles to the projection \( v \) of the velocity of the ion on that plane, and equal to \( evr^{-2} \). For vibrations whose wavelength in free ether is very great compared with the dimensions of the molecular orbit, if we interpret magnetic force as velocity of the ether, the vibration-path of a point attached to the ether, and close to the vibrator, will be in the plane transverse to \( r \), and similar to the projection of the orbit of the electron on that plane when turned round through a right angle. This was described by Lorentz since 1892, and was used to explain the Zeeman effect. If the condition of wavelength very large compared with molecular magnitude were not satisfied, each spectral line would be accompanied, more or less, by its system of harmonics.

On the basis of this analysis Larmor says that it is verified that the free periods of the radiation are those of the system of ions.

It might appear that every steady orbital motion must rapidly lose its energy by radiation, so that the permanency of atoms is not held. Larmor discusses this problem. He says — "We might assume that non-radiating
vibrations consisted of stationary waves reflected backwards and forwards between two vibrating molecules, or between two ions in the same molecule; but even that would not be satisfactory. As a matter of fact, however, no explanation of this kind is needed." He gives the following reasoning. The effective electric inertia of an ion e is \((2/3)e^2a^{-1}\), where \(a\) is the radius of the ion supposed spherical.\(^{54}\) The rate at which it loses energy by radiation is proportional to \(e^2\) and involves its motion, but does not depend on \(a\) at all. The kinetic reaction to change of its velocity which is connected with the loss of energy by radiation, can thus be made negligible in comparison with the kinetic reaction arising from its inertia. Larmor considers so from the fact that the energy of the ethereal motion carried along by the moving ion depends on the first term in \(H\), involving \(r^{-2}\), and the radiated energy depends on the second term, involving \(r^{-1}\).

He calculates then the electric force and integrates the Poynting vector over an infinite sphere, to derive the rate of loss of energy by radiation \((2/3)e^2c^{-1}\dot{u}^2\) (in electromagnetic units). The energy emitted by radiation in the process of getting up a velocity \(u\) of the ion from rest is given by \((2/3)e^2c^{-1}\int \dot{u}^2 dt\).

The rate of loss of energy by radiation is thus \((2/3)e^2c^{-1}\times (\text{acceleration})^2\), while the kinetic energy belonging to the ion is \((2/3)e^2a^{-1}u^2\). Accordingly, Larmor considers that the loss of energy by radiation from an undisturbed vibrating molecule would not be sensible compared with its whole intrinsic kinetic energy, when the velocities of the ions are not of the order of magnitude of that of radiation. He considers also that for higher velocities the importance of the radiation is counteracted by the increase of the inertia coefficient.

The magnetic vibration excited by a moving ion is treated above only for the case in which \(r\) is small compared with the wavelength. Further away from the ion the variation of the magnetic force with distance is given by \(eu/r^2 + e\dot{u}/cr\) as the equation for \(H\) shows. Therefore, the vibration-curve of the radiation proper is similar to the projection of the hodograph of the orbit of the ion on the wave-front, instead of the projection of the orbit itself.

From these considerations Larmor arrives at the following conclusion about the permanence of atoms in particular — "When the steady orbital motions in a molecule are so constituted that the vector sum of the accelerations of

\(^{54}\)Larmor cites Phil. Trans. A, 1894, p. 812. He points out that this inertia changes with the motion, citing Phil. Trans. A, 1895, p. 718 for the computation in the case of the velocity of the ion considerably compared with that of radiation.
all its ions or electrons is constantly null, there will be no radiation, or very little, from it, and therefore this steady motion will be permanent. But this is just the condition which holds good so long as the molecule is free from extraneous disturbance."

4. Model of the Atom without Nucleus

§1. Study of cathode rays by J. J. Thomson

Hertz\(^{55}\) examined whether cathode rays were deflected by the electrostatic field produced in a discharge tube by a condenser or not, with negative results. He concluded thereby that cathode rays were not caused by electrically charged particles but by phenomena of ether. He failed because in his experiments the degree of vacuum was not high, so that cathode rays were not effectively affected by the electrostatic field on account of the ionization of remnant gases.

Lenard\(^{56}\) put a window closed by a metallic film on a discharge tube, and succeeded in extracting cathode rays through it. For this reason cathode rays were called Lenard rays. Because Lenard rays had an electric charge, cathode rays were concluded not to be molecules. Lenard supposed, however, that they were not material particles but a series of waves in ether.

J. J. Thomson made also experiments of deflecting cathode rays by an electrostatic field, but did not succeeded in it in his early experiments.\(^{57}\) With the degree of vacuum of the discharge tube made higher, cathode rays were deflected by means of a charged condenser, and thus the measurement of \(e/m\) of a free electron in cathode rays was achieved for the first time to verify the existence of the electron.\(^{58}\)

In the introductory remarks of his paper cited now, Thomson says — "The most diverse opinions are held as to these rays; according to the almost unanimous opinion of German physicists they are due to some process in the aether to which — inasmuch as in a uniform magnetic field their course is circular and not rectilinear — no phenomenon hitherto observed is analogous: another view of these rays is that, so far from being wholly material, they are in fact

\(^{55}\)H. R. Hertz, Wied. Ann. 19 (1883), 782; 45 (1892), 28.
\(^{56}\)P. Lenard, Wied. Ann. 51 (1894), 225; 52 (1894), 23.
\(^{57}\)J. J. Thomson, Camb. Phil. Soc. 9 (1897) and Electrician, May 21, 1897, as cited by Thomson in Ref. 58).
\(^{58}\)J. J. Thomson, Phil. Mag. 44 (1897), 293.
wholly material, and that they mark the paths of particles of matter charged with negative electricity. It would seem at first sight that it ought not to be difficult to discriminate between views so different, yet experience shows that this is not the case, as amongst the physicists who have most deeply studied the subject can found supporters of their theory."

After having described so the former views of cathode rays, Thomson presents an epistemological opinion about the superiority between the two views from a materialistic standpoint as follows — "The electrified-particle theory has for the purpose of research a great advantage over the aetherial theory, since it is definite and its consequences can be predicted; with the aetherial theory it is impossible to predict what will happen under any given circumstances, as on this theory we are dealing with hitherto unobserved phenomena in the aether, of those laws we are ignorant." Thomson, taking a definite viewpoint, thus goes into experiments from the standpoint of examining a substantialistic theory — "The following experiments were made to test some of the consequences of the electrified-theory."

If the cathode rays are negatively charged particles, then when they enter an enclosure they ought to carry into it a charge of negative electricity, which could be detected with an electroscope. Such an experiment was made for the first time by Perrin, who placed in front of plane cathode two coaxial metallic cylinders: the outer of them was connected with the earth, the inner with an electroscope. Perrin found that when the rays passed into the inner cylinder the electroscope received a charge of negative electricity, while no charge went to the electroscope when the rays were deflected by a magnet so as no longer to enter the inner cylinder.

Thomson points out however that Perrin's experiment is open to the objection that it does not prove that the cause of the electrification in the electroscope has anything to do with the cathode rays — "Now the supporters of the aetherial theory do not deny that electrified particles are shot off from the cathode; they deny, however, that these charged have any more to do with the cathode rays than a rifle-ball has with the flash when a rifle is fired." Thomson therefore repeated Perrin's experiment in a form which is not open to this objection. In other words, the cathode rays were made not to fall upon the cylinders unless they deflected by a magnet. The path of the cathode rays was traced by the phosphorescence on the glass. When the rays were bent by a magnet so as to fall on the cylinders there was a great response of the electrometer. Thus this experiment shows that however we twist and deflect the
cathode rays by magnetic forces, the negative electrification follows the same path as the rays, and that this negative electrification is inseparably connected with the cathode rays.

In the next place observation was made of the deflection of cathode rays by an electrostatic field. Thomson says — “An objection very generally urged against the view that the cathode rays are negatively charged particles, is that hitherto no deflection of the rays has been observed under a small electrostatic force, and though the rays are deflected when they pass near electrodes connected with sources of large differences of potential, such as induction-coils or electrical machines, the deflection in this case is regarded by the supporters of the aetherial theory as due to the discharge passing between the electrodes, and not primarily to the electrostatic field.” Hertz found that the cathode rays were not deflected when the plates were connected with a battery. On repeating Hertz’s experiment, Thomson showed that the absence of deflection was due to the conductivity conferred on the rarefied gas by the cathode rays. The conductivity diminished very rapidly as the exhaustion increased. The rays were deflected when the two plates were connected with a battery of small storage-cells with the potential difference of a small as 2 volts. It was only when the vacuum was a good one that the deflection took place.

Then, Thomson made experiments on the deflection of the cathode rays exerted by a magnetic field in different gases. The magnetic field was generated by placing “a bell-jar between two large parallel coils arranged as a Helmholtz galvanometer”.

Thomson states — “As the cathode rays carry a charge of negative electricity, are deflected by an electrostatic force as if they were negatively electrified, and are acted on by a magnetic force in just the way in which this force would act on a negatively electrified body moving along the path of these rays, I can see no escape from the conclusion that they are charges of negative electricity carried by particles of matter.” Having argued in this way that this conclusion was undeniable, Thomson “made a series of measurements of the ratio of the mass of these particles to the charge carried by it”, “to throw some light on the point” that “the question next arises, What are these particles? are they atoms, or molecules, or matter in a still finer state of subdivision?”, arriving at the solution.

Let \( m \) be the mass of each of the particles, \( e \) the charge carried by it, and \( N \) the number of particles passing across any section of the beam in a given
time. Then $Q$, the quantity of electricity carried by these particles is given by the equation

$$Ne = Q.$$  \hfill (81)

We can measure $Q$ if we receive the cathode rays in the inside of a vessel connected with an electrometer. When these rays strike against a solid body, the temperature of the solid body is raised, the kinetic energy of the moving particles being converted into heat. If we suppose that all this energy is converted into heat, and if we measure the increase in the temperature of a body of known thermal capacity, we can determine $W$, the kinetic energy of the particles

$$Nmu^2 = 2W, \quad \hfill (82)$$

$u$ being the velocity of the particles.

If $\rho$ is the radius of curvature of the path of these rays in a uniform magnetic field $H$, we have

$$\frac{mu}{\rho} = H\rho \equiv I. \quad \hfill (83)$$

From these equations we get (in electromagnetic units)

$$\frac{mu^2}{2e} = \frac{W}{Q}, \quad \hfill (84)$$

$$u = \frac{2W}{QI}, \quad \hfill (85)$$

$$\frac{m}{e} = \frac{I^2Q}{2W}. \quad \hfill (86)$$

Thus, if we know the values of $Q$, $W$ and $I$, we can deduce the values of $u$ and $m/e$.

To measure these quantities, Thomson used tubes of three different types. The measurement of the temperature was made by means of a thermoelectric couple of iron and copper. The radius of curvature $\rho$ was determined from the measurement of the position at which the cathode rays after having passed through the slit struck against the tube. The measurement of $Q$ was complicated by the leakage of the charge caused by the cathode rays making the gas a conductor. For this reason, the inner cylinder to receive the cathode
rays was connected with a large capacity of 1.5 microfarad, and the rays were only kept on for a short time of about 1 or 2 seconds, so that the alternation in potential of the inner cylinder was not large, ranging from about 0.5 to 5 volts. Another reason why the duration of the rays was limited was to avoid the loss of heat from the thermoelectro-junction by conduction through the wires. The rise in temperature of the junction was of the order $2^\circ$C. A series of experiments showed that with the same tube and the same gaseous pressure $Q$ and $W$ were proportional to each other, when the rays were not kept on too long. The values of the quantities were different respectively for the tubes of different types.

The value of $m/e$ is thus found to be independent of the nature of the gas. For the first tube this value is about $0.4 \times 10^{-7}$ ($e/m = 2.2 \times 10^7$), for the second about $0.5 \times 10^{-7}$ ($e/m = 2.0 \times 10^7$), and for the third $0.9 \times 10^{-7}$ ($e/m = 1.1 \times 10^7$). Thomson considers the value for the second tube best because this tube has cylinders of small opening.

Thomson took also another method of measuring $m/e$ and $u$. It is based upon the deflection of the cathode rays in an electrostatic field, in addition to that due to a magnetic field. This method is the one that is applied nowadays to mass-spectrometers. Let the space passed over by the rays under a uniform electric field $F$ be $l$. The time taken for the rays to traverse this space is $l/u$, and the velocity in the direction of $F$ is $(Fe/m)(l/u)$, so that $\theta$, the angle through which they are deflected when they leave the electric field, is given by

$$\theta = \frac{Fe}{m} \frac{l}{u^2}.$$  \hspace{1cm} (87)

If the rays are acted on by a magnetic field $H$ at right angle to the rays, and extending across the distance $l$, the velocity at right angle to the original path of the rays is $(Heu/m)(l/u)$, so that $\phi$, the angle through which the rays are deflected when they leave the magnetic field is given by

$$\phi = \frac{He}{m} \frac{l}{u}.$$  \hspace{1cm} (88)

From these equations we get

$$u = \frac{\phi F}{\theta H}.$$  \hspace{1cm} (89)
and
\[ \frac{m}{e} = \frac{H^2 \theta l}{F \phi^2}. \quad (90) \]

In Thomson’s experiments \( H \) has adjusted so that \( \phi = \theta \). In this case we have
\[ u = \frac{F}{H}, \quad (89') \]
and
\[ \frac{m}{e} = \frac{H^2 l}{F \phi}. \quad (90') \]

The apparatus used for this purpose was the same as that previously used for the deflection due to an electrostatic field. The magnetic field was produced by placing outside the tube two coils, whose diameter was equal to the length of the plates producing the electric field. The coils were placed so that they covered the place occupied by the plates, and the distance between the coils was equal to the radius of either. The value of \( m/e \) obtained with this method was \( 1.1 \sim 1.5 \times 10^{-7} \) \( (e/m = 0.7 \sim 0.9 \times 10^7) \).

About this Thomson says — “From these determinations we see that the value of \( m/e \) is independent of the nature of the gas, and that its value \( 10^{-7} \) is very small compared with the value \( 10^{-4} \), which is the smallest value of this quantity previously known, and which is the value for the hydrogen ion in electrolysis.”

Then, Thomson enters into the problem that “the smallness of \( m/e \) may be due to the smallness of \( m \) or the largeness of \( e \), or to a combination of these two”. “That the carriers of the charges in the cathode rays are small compared with ordinary molecules is known, I think, by Lenard’s results as to the rate at which the brightness of the phosphorescence produced by these rays diminishes with the length of path traveled by the ray”, Thomson says. If this phosphorescence is regarded as due to the impact of the charged particles, the distance through which the rays must travel before the phosphorescence fades to a given fraction (say \( 1/e \), where \( e = 2.71 \)) of its original intensity, will be some moderate multiple of the mean free path. Lenard found that this distance depended solely upon the density of the medium, and not upon its chemical nature or physical state. In air at atmospheric pressure the distance was about half a centimeter, and was considered to be comparable with the mean free path of the carriers through air at atmospheric pressure. Thomson concludes,
however, that "the mean free path of the molecules of air is a quantity of quite a different order", and "the carrier, then, must be small compared with ordinary molecules".

Thomson points out the following two points as the fundamentals about these carriers:

1. these carriers are the same whatever the gas through which the discharge passes.
2. the mean free paths depend upon nothing but the density of the medium traversed by these rays.

From the first point Thomson concludes that the mass of the carriers is an electromagnetic one, saying that "the mass concerned is the quasi mass which a charged body possesses in virtue of the electric field set up in its neighbourhood". He points out that the moving body involves the production of a varying electric field, and, therefore, of a certain amount of energy which is proportional to the square of the velocity.

\[ \textbf{§2. Thomson's atomic model in terms of corpuscles} \]

From the above results Thomson presents a theory of atomic constitution in which the particles concerned are supposed to make up atoms. He says — "The explanation which seems to me to account in the most simple and straightforward manner for the facts is founded on a view of the constitution of the chemical elements which has been favourably entertained by many chemists: this view is that the atoms of the different chemical elements are different aggregations of atoms of the same kind." In the hypothesis enunciated by Prout, the atoms of the different elements were hydrogen atoms. Thomson affirms that in this precise form the hypothesis is not tenable, but if some unknown primordial substance X is substituted for hydrogen, there is nothing known which is inconsistent with this hypothesis. He says in addition that this hypothesis is the one that has been recently supported by Sir Norman Lockyer for reasons derived from the study of the stellar spectra. Thomson supposes thus that X is the particle of cathode rays.

According to this theory Thomson explains the generation of cathode rays. In the very intense electric field in the neighbourhood of the cathode, the molecules of the gas are dissociated and are split up, not into the ordinary chemical atoms, but into these primordial atoms, "which we shall for brevity
call corpuscles. If these corpuscles are charged with electricity and projected from the cathode by the electric field, they would behave exactly like the cathode rays. They would evidently give a value of $m/e$ which is independent of the nature of the gas and its pressure. And again, the mean free paths of these corpuscles would depend solely upon the density of the medium which they pass. For, the molecules of the medium are composed of a number of such corpuscles separated by considerable spaces, and the collision between a single corpuscle and the molecule will not be between the corpuscles and the molecule as a whole, but between this corpuscle and the individual corpuscles which form the molecule. Thus, the number of collisions the particle makes as it moves through a crowd of these molecules will be proportional, not to the number of the molecules in the crowd, but to the number of the individual corpuscles. The mean free path is inversely proportional to the number of collisions in unit time, and so is inversely proportional to the number of corpuscles in unit volume. As these corpuscles are all of the same mass, the number of corpuscles in unit volume will be proportional to the mass of unit volume, that is, the mean free path will be inversely proportional to the density of the gas. It is seen, too, that so long as the distance between neighbouring corpuscles is large compared with the linear dimensions of a corpuscle, the mean free path will be independent of the way they are arranged, and therefore independent of its chemical nature and physical state. Thomson explains in this way that the result from Lenard’s very remarkable measurements of the absorption of the cathode rays by various media, must be a property possessed by the carriers of the charges in the cathode rays.

Thus Thomson arrives at the view that “we have in the cathode rays matter in a new state, a state in which the subdivision of matter is carried very much further than in ordinary gaseous state: a state in which all matter — that is, matter derived from different sources such as hydrogen, oxygen, &c. — is of one and the same kind; this matter being the substance from which all the chemical elements are built up”. This view means to suppose that matter consists only of electrons and the mass of matter is due to the mass of the electrons. This supposition is recognized a little later by Larmor, as we have seen in the last section on the Zeeman effect.

So, the cathode rays are supposed to be produced by the projection of the fundamental elements of matter. However, the quantity of matter produced by means of the dissociation at the cathode is so small as to almost to preclude the possibility of any direct chemical investigation of its properties.
Thomson calculates that the apparatus he used would, if kept going uninterruptedly night and day for a year, produce only about one three-millionth part of a gramme of this substance. In other words, though the new elementary atoms consisting of the cathode rays are thought of being freed from molecules, their amount is so small that they could not be studied chemically.

Thomson thinks, in his paper under consideration, that the smallness of the value of \( \frac{m}{e} \) is due to the largeness of \( e \) as well as the smallness of \( m \). To him "there seems to be some evidence that the charges carried by the corpuscles in the atom are large compared with those carried by the ions of an electrolyte". As the reason for this he mentions that the specific inductive capacity of a compound is very approximately given by adding those of constituent atoms, and hence each atom would have a electrical moment large compared with that of the molecule. Thomson pictures, in the molecule of HCl, for example, the components of the hydrogen atom as held together by a great number of tubes of electrostatic force, the component of the chlorine atom being similarly held together, while only one stray tube binds the hydrogen atom to the chlorine atom. He concludes that, for the electrical moment of the atom to be large compared with that of molecule, the charge on the corpuscles would have to be very large compared with those on ions. It is not understandable, however, that no consideration is given by Thomson of the number of the corpuscles.

Finally Thomson enters into the problem of atomic structure. He says that, if we regard the chemical atom as an aggregate of a number of primordial atoms, the problem of finding the configuration of stable equilibrium for a number of equal particles acting on each other according to some law of force, is of great interest in connection with the relation between the properties of an element and its atomic weight. About the law of force concerned he mentions that of Boscovich, where the force between them is a repulsion when they are separated by less than a certain critical distance, and an attraction when they are separated by a greater distance, or even the simpler case of a number of mutually repellent particles held together by a central force. The equations to determine the stability of such a collection of particles increase so rapidly in complexity with the number of particles, that a general mathematical investigation is scarcely possible. Thomson tries, however, to obtain the general laws which govern such configurations by the use of models. As the simplest of them he takes Mayer’s floating magnets. In this model the magnets arrange themselves in equilibrium under their mutual repulsion and a central attraction caused by the pole of a large magnet placed above the floating magnets.
Thomson says that to him "a study of the forms taken by these magnets seems to be suggestive in relation to the periodic law". The model of floating magnets remained one of his dominant ideas until later times. The fundamental point of the atomic model in terms of corpuscles consisted fairly in explaining the chemical properties of atoms by means of the number of corpuscles and the periodicity of their configurations. It may be said that the fundamental point of the aim of the present-day theory of atomic structure was here established.

Mayer's arrangements of magnets are as follows

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1.5</td>
<td>2.6</td>
<td>3.7</td>
<td>4.8</td>
<td>5.9</td>
</tr>
<tr>
<td>2.</td>
<td>1.6</td>
<td>2.7</td>
<td>3.7</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>1.7</td>
<td>1.5.9</td>
<td>2.7.10</td>
<td>3.7.10</td>
<td>4.8.12</td>
</tr>
<tr>
<td></td>
<td>1.6.9</td>
<td>2.7.11</td>
<td>3.7.11</td>
<td>4.8.13</td>
<td>5.9.13</td>
</tr>
<tr>
<td></td>
<td>1.6.10</td>
<td>2.8.10</td>
<td>3.8.10</td>
<td>4.9.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.11</td>
<td></td>
<td>3.8.11</td>
<td>4.9.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.8.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.8.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5.9.12</td>
<td>2.7.10.15</td>
<td>3.7.12.13</td>
<td>4.9.13.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5.9.13</td>
<td>2.7.12.14</td>
<td>3.7.12.14</td>
<td>4.9.13.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.9.12</td>
<td>2.7.12.14</td>
<td>3.7.13.14</td>
<td>4.9.14.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.10.12</td>
<td></td>
<td>3.7.13.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.10.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.11.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.11.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.11.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6.11.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.7.12.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here, for example, 1.6.10.12 means an arrangement with one magnet in the middle, then a ring of six, then a ring of ten, and a ring of twelve outside.

Now suppose that a certain property is associated with two magnets forming a group by themselves. This property occurs in the first place in the arrangement 2, and then in those written below it. In other words, we should
have this property with 2 magnets, again with 8 and 9, again with 19 and 20, and again with 34, 35, and so on. If the system of magnets is regarded as a model of an atom, the number of magnets is proportional to the atomic weight. The above property should occur then in elements of atomic weight 2, (8, 9), 19, 20 (34, 35). One obtains thus something quite analogous to the periodic law, the first series corresponding to the arrangement of the magnets in a single group, the second series to the arrangement in two groups, the third series in three groups, and so on.

Thomson put the atomic weight equal to the number of corpuscles, because he supposed that the mass of every atom was made up solely of those of corpuscles, or, that every atom consisted only of corpuscles. It was not correct, however, to relate the atomic weight to the number of corpuscles.

§3. On the electric charge of electron

The measurement of \( e/m \) established the substantiality of the electron. Its value of cathode-ray particles was found, however, very much larger than that of positively charged particles, as was described in Thomson's words. Thomson considered at first this was due not only to the smallness of \( m \) but also to the difference in \( e \). From about that time the study of ionization of gas was advanced by J. J. Thomson, E. Rutherford, J. S. Jownsent, C. T. R. Wilson and others, mainly persons in Cavendish Laboratory. Particularly in Wilson's method use was made of the formation of drops of water by the ions generated in gas in a state of super saturation of vapor of water as nuclei. This study gave the basis for a means of observation which plays an important role nowadays as the Wilson cloud chamber in the study of atomic nuclei and cosmic rays. Wilson attained the formation of drops of water by rapidly expanding gas adiabatically. He irradiated X-rays on gas to cause ionization and examined clouds so produced. In a similar way, the process of ionization due to \( \beta \)- or \( \alpha \)-rays emitted by radioactive elements was also studied.

It was found that there exists the minimum quantity of electric charges, and the positive and negative minimum charges are identical in their quantities. This confirmed the electric charge of the electron. The history of the measurement of the minimum electric charge is given in detail in Millikan's book,\(^59\) to which the reader is referred.

The problem of the existence of the minimum electric charge began with Faraday’s study of electrolysis. This showed that, when an electric current passes through an electrolytic solution, the amount of negative electric charge gathered on the positive plate as well as that of positive one gathered on the negative plate is determined by the number of atoms gathered on an electric plate. An atom of every monovalent element gives, irrespective of its kind, the same amount of electric charge to an electric plate as a hydrogen atom, that of every divalent or trivalent element giving two or three times this amount, respectively. Electric charges appearing in the process of electrolysis were found always to be integral multiples of that carried by a hydrogen atom, no fraction of it being observed. A definite indication of the “atom of electricity” was given by Helmholtz in his Faraday Lecture.\(^6\)

About the minimum electric charge, a definite value was obtained for the first time by J. J. Thomson\(^6\) in 1898. He irradiated gases under a certain electric voltage with X-rays, and measured the electric current passing through them to determine the values of \(n e\), where \(n\) is the number of ions per unit volume of a gas, \(e\) the electric charge of an ion, and \(u\) the average velocity of positive as well as negative ions under the imposed electric voltage. Values of \(u\) were measured by Rutherford\(^6\) in 1897 for various kinds of gas. Thomson used them to deduce the values of \(n e\). The measurement of \(n\) would then determine \(e\).

To the measurement of \(n\) Thomson applied the method of C. T. R. Wilson\(^6\) we have mentioned above. In this case, if every ion is assumed to make the nucleus of a drop of water, the number of the drops of water formed in a gas of unit volume, and hence the number of the ions produced in it, can be known from the size of the drops of water and the quantity of the water settled in it. The first object of the experiment under consideration was therefore to know the size of the drops of water, which would determine \(n\), and the second to know the quantity of electric charge, which would determine \(n e\). From their results \(e\) would be determined.

Thomson made use of optical and other methods to know the size of the drops of water, but obtained inaccurate results. He thus applied finally Stokes’


\(^{61}\)J. J. Thomson, Phil. Mag. 46 (1898), 528.

\(^{62}\)E. Rutherford, Phil. Mag. 44 (1897), 422.

\(^{63}\)C. T. R. Wilson, Phil. Trans. A. (1897), 265; (1899), 440.
method, in which the sinking speed \( w \) of the drops of water is measured to determine their radius \( a \) according to the formula

\[
w = \frac{(2/9)ga^2}{\mu},
\]  

(91)

where \( \mu \) is the coefficient of viscosity of the gas and \( g \) is the gravity constant.

The sinking velocity was measured by observing the upper limit of the cloud of the drops of water. The value of \( e \) obtained in this way was

\[e = 6.5 \times 10^{-10} \text{ e.s.u.}.
\]

This value is larger than the present-day value of \( e \).\(^{64}\) This is due to the calculation made on the assumption, as was pointed out later by Thomson,\(^ {65}\) that all the positive ions would form drops of water, whereas positive ions do not necessarily form drops of water at expansion ratios of the order of magnitude of that used in Thomson's experiment.

It should be noted that Thomson determined in his paper under consideration the Avogadro number. From the measurement of electrolysis it was known that

\[Ne = 129 \times 10^8,
\]

where \( N \) is the number of molecules per cm\(^3\) of a gas in the standard state and \( e \) the electric charge of a hydrogen ion (e.s.u.).\(^ {66}\) Combining this with the value of \( e \) given above, Thomson obtained

\[N = 20 \times 10^{18}.
\]

About this result Thomson wrote — “Though the measurements of the coefficients of viscosity of other gases give in general higher values of \( N \), yet the agreement between the value of \( N \) deduced from these experiments and the value of \( N \) got by the Kinetic Theory of Gases by viscosity experiments is sufficient to show that theory is consistent with the value we have found for \( e \) being equal to, or at any rate of the same order as, the charge carried by the hydrogen ion in electrolysis.”

He then wrote about Lorentz's result — “In connection with this result it is interesting to find that Professor H. A. Lorentz has shown that the charge

\(^{64}\)e = 4.804 \times 10^{-10} \text{ e.s.u.}
\(^{65}\)J. J. Thomson, Phil. Mag. 5 (1903), 346.
\(^{66}\)Thomson used Richarz's measurement made in 1891.
on the ions whose motion causes those lines in the spectrum which are affected by the Zeeman effect is of the same order as the charge on the hydrogen ion in electrolysis." Thus, the identity was shown of the ion, the substance supposed theoretically from the Zeeman effect, and the corpuscle, the substance deduced from experiments on cathode rays and on ionization of gases.

Subsequently to Thomson's work the measurement of the minimum electric charge was advanced by means of various methods. In 1903 J. J. Thomson took into consideration the phenomenon fund by C. T. R. Wilson, that drops of water do not attach to positive ions but only to negative ions, when the expansion ratio of a gas in a state of over-saturation of vapor of water is less than a certain value, namely when the degree of over-saturation is low, and made his experiment again, getting a better value.

Also in 1903 H. A. Wilson\(^{67}\) chose an expansion ratio such as to make the cloud attach to negative ions, that is, such as to make the droplets produced carry only negative electric charges, and determined first the size of a droplet by observing its rate of fall in the manner mentioned above. He then put a positively charged plate above the cloud to pull up it. When the electric field is adjusted so as to balance with the gravitational field, the droplets stay at rest in the air. If \(X\) is the strength of the electric field, the electric force acting on a droplet is \(Xe\), and should be equal to its weight. Thus, by knowing \(X\), the value of \(e\) can be determined. Millikan used later oil-droplets instead of water-droplets and determined accurately the value of \(e\) by means of this method after detailed corrections.

Townsend determined the value of \(e\) by means of the diffusion of ions through gas. He made this fundamental study on ions in gas around 1900.\(^{68}\) According to him, let us treat the diffusion of ions in gas between two horizontal planes. In the steady state the number of ions in every thin layer between the planes is assumed to be constant. If \(x\) is the distance to a layer from the lower plane, \(n\) the number of ions of a single sign per unit volume in the layer, and \(D\) the diffusion coefficient, the number of the ions passing downward through the layer per unit area and per second is given by

\[
D \frac{dn}{dx}.
\]

\(^{67}\)H. A. Wilson, Phil. Mag. 5 (1903), 429.

\(^{68}\)J. S. Townsend, Phil. Trans. A193 (1899), 129.
The average velocity of the going down ions is hence
\[
\frac{D}{n} \frac{dn}{dx}.
\]

The force making the ions move is due to the change in the partial pressure. If the partial pressure is \( p \), the force acting on the ions in a unit volume is \( dp/dx \), and the average force acting on an ion is \( (dp/dx)/n \). In the case of a known force acting on an ion, the velocity gained by the ion can be known from measurements by means of the study of Rutherford and others. It is possible therefore to know the velocity gained by an ion in an electrostatic field. According to Rutherford and others this velocity is proportional to the force acting on the ion. Let \( A \) be the velocity in the case of the electrostatic field \( X \). The force acting on the ion is \( Xe \), and the velocity per unit force is \( A/Xe \). The velocity in the case of the force \( (dp/dx)/n \) is therefore given by
\[
\frac{1}{n} \frac{dp}{dx} \frac{A}{Xe}.
\]

This must be equal to \( (D/n)(dn/dx) \), so that
\[
\frac{dp}{dx} \frac{A}{Xe} = D \frac{dn}{dx}. \tag{92}
\]

Now, if the ions are supposed to obey the law of perfect gas, \( p \) is proportional to \( n \). The ratio between them at a given temperature is identical for all kinds of gases. If \( N \) is the Avogadro constant (the number of molecules per cubic centimeter at the atmospheric pressure \( P \)), one has
\[
\frac{p}{P} = \frac{n}{N}. \tag{93}
\]

Eq. (92) becomes accordingly
\[
\frac{PA}{XD} = Ne, \tag{94}
\]

so that \( Ne \) can be known if \( D \) and \( A \) are known. In this way Townsend found that \( Ne \) takes one and the same value for air, hydrogen gas, oxygen gas and carbonic acid gas, the average value being \( Ne = 1.24 \times 10^{10} \). If \( E \) is the electric charge of a hydrogen ion, one has \( NE = 1.22 \times 10^{10} \). It results that
\[
e = E, \tag{95}
\]
or, in other words, that the electric charge of a gas ion is equal to the electric charge carried by a hydrogen ion in electrolysis.

About the objection that the electric charge measured in these experiments would be that of molecules and would not be that of electrons, Thomson made then a careful experiment. He did not directly ionized gas but irradiated a surface of metal to make electrons to be ejected from it, and measured ions which were formed by the attachment of the electrons to gas molecules.

Let us mention some words here about the naming of electron. The name "electron" was used first by G. J. Stoney in 1891, in the meaning of the natural unit of electricity. He supposed electrons to be "the charges of electricity which are associated with chemistry bonds, and which, so long as they are undisguised, are acted on by the disturbance perpetually going on in the surrounding aether", and treated in particular chemistry and electrolysis. To them the present-day term ions would rather correspond. Contrarily Lorentz used in the same age the word ions for electrons, as we have mentioned above. Thomson used in particular the word corpuscles for electrons. Since Drude (1900), only the carriers of electricity in electrolysis were called ions. Kelvin used, as we shall describe in the following, "electrions", namely electro-ions for electrons.

§4. **Kelvin's model of the atom**

In 1902, Kelvin made clear an idea of model of atoms without nucleus, in the direction similar to J. J. Thomson's idea in his argument by analogy to floating magnets, and discussed the nature of various kinds of matter by means of it. Kelvin gave the title "Aepinus Atomized" to his paper. F. U. T. Aepinus was a professor of physics at the university of Petersburg, who adopted B. Franklin's one-fluid theory of electricity. Kelvin followed Aepinus's one-fluid theory, but supposed that this fluid of electricity would consist of "electrions" as the atoms of electricity, to set up a theory of the nature of matter. Kelvin's idea

---

70) G. J. Stoney, Phil. Mag. 40 (1895), 372.
71) Lord Kelvin, Phil. Mag. 3 (1903), 257.
72) Aepinus considered that an excess or lack of the fluid of electricity would cause a positive or negative electric charge, respectively. He considered also that parts of matter lacking the fluid of electricity would repulse each other, parts of the fluid of electricity would repulse each other, and a part of the fluid of electricity and a part of matter lacking it would attract each other.
was suggested first in 1897, a little before J. J. Thomson's determination of
the value of $e/m$ in the study of cathode rays, in his paper\(^73\) which has the
title "Contact Electricity and Electrolysis".

In Ref. 73) Kelvin says — "Varley's fundamental discovery of the cathode
torrent,\(^74\) splendidly confirmed and extended by Crookes, seems to me to ne-
cessitate the conclusion that resinous electricity, not vitreous, is the electric fluid, if we are to have a one-fluid theory of electricity.\(^75\) Mathematical rea-
sons, to which I can only refer without explanation at present, prove that if
resinous electricity is a continuous homogeneous liquid, it must, in order to re-
produce the phenomena of contact-electricity which you have seen this evening,
be endowed with a cohesional quality \ldots. It is just conceivable, though it does
not at present seem to me very provable, that this idea may deserve care-
ful consideration. I leave it, however, for the present, and prefer to consider
an atomic theory of electricity foreseen as worthy of thought by Faraday and
Clerk Maxwell, very definitely proposed by Helmholtz in his last lecture to
the Royal Institution, and largely accepted by present-day theoretical workers
and teachers."

Kelvin then discusses the problem of naming the atom of electricity —
"Indeed, Faraday's law of electro-chemical equivalence seems to necessitate
something atomic in electricity, and to justify the very modern name electron. The older, and at present even more popular, name ion given sixty years ago
by Faraday, suggests a convenient modification of it, electrion, to denote an
atom of resinous electricity. And, now, adopting the essentials of Aepinus' theory, and dealing with it according to the doctrine of Father Boscovich, each
atom of ponderable matter is an electron of vitreous electricity,\(^76\) which, with
a neutralizing electrion of resinous electricity close to it, produces a resulting
force on every distant electron and electrion which varies inversely as the cube
of the distance, and is in the direction determined according to the well-known
requisite application of the parallelogram of forces."

In his paper in 1902, Ref. 71), the above consideration was developed
to a theory of model of the atom without nucleus, in which electrions were

\(^{73}\) Lord Kelvin, Nature 56 (1897), 84.
\(^{74}\) C. F. Varley (1828-1883). [Present author's note]
\(^{75}\) Resinous electricity means negative electricity, and viterous electricity means positive electricity. These two kinds of electricity were confirmed and named so by Dufay (1733). They were given the terms positive and negative by Franklin (1747). [Present author's note]
\(^{76}\) Note that Kelvin applies the name electron to the atom of positive electricity.
supposed to be much smaller than atoms of ponderable matter. Electrons were assumed to path through freely any space out of or in existence of atoms. Similarly to Aepinus's theory, electrons were supposed to repulse each other, atoms without electron to repulse each other, and every electron and every atom without electron to attract each other.

In the paper under consideration, Kelvin supposes that, in accordance with Cavendish and Coulomb's law of inverse square of distance, the atoms which he assumes to be spherical, repel other atoms outside them with forces inversely as the squares of distances between centers. About electrons he considers that they "no doubt occupy finite spaces, although at present we are dealing with them as if they were mere mathematical points". He assumes also that every atom attracts every electron outside it with a force inversely as the square of the distance between them.

"As a tentative hypothesis", he assumes "for simplicity" that the attraction experienced by an electron approaching an atom varies according to the inverse square of the distance from the center, as long as the electron is outside, has no abrupt change when the electron enters the atom, and decreases as the distance from the center when the electron is within the spherical boundary of the atom.\(^{77}\) This is just as it would be if there were a uniform distribution through the atom of an ideal electric substance. But Kelvin considers that the corresponding supposition cannot be made for the mutual force between two overlapping atoms, for the reason that it must be kept free to add any force that may be convenient for the explanation of electric, elastic and chemical properties of matter.

The neutralizing number of electrons for atom or group of atoms has exactly the same quantity of electricity of one kind as the atom or group of atoms has of electricity of the opposite kind. The number for any single atom may be an integer, but needs not be the same for all atoms. Kelvin considers that it is possible that the differences of quality of the atoms of different substances may be partially due to the numbers of their electrons being different, but it is also possible that the differences of quality are to be wholly explained in merely Boscovichian fashion by differences in the laws of forces between atoms, and may not imply any difference in the numbers of electrons.

In addition he keeps in view another possibility that the neutralization number for an atom may not be any integer. For example, he says, the molecule

\(^{77}\)This agrees with the force inside an atom which Lorentz assumed in his study of the Zeeman effect, although no mention is made of it in Kelvin's paper.
of a diatomic gas might conceivably have some odd number of electrions so that the single atoms must be positively or negatively electrified and cannot be neutral.

Kelvin says that in the mathematical theory of electricity in equilibrium, it is a matter of perfect indifference which of the opposite electric manifestations we call positive and which negative, but he points out that the great differences in the disruptive and luminous effects, when the forces are too strong for electric equilibrium, "show physical properties not touched by the mathematical theory". Thus, in a substantialistic way independently of the mathematical theory of electricity, he arrives at assuming a fluid of negative electricity for cathode rays, and setting up a theory of matter in which a void atom acts simply as a little globe of atomic substance, possessing positive electricity uniformly distributed through it.

Kelvin considers that any atom having just enough electrions to neutralize it exerts no electric force, at large distances compared with its radius, because of the inverse square law of electric force. But, he thinks, even if every neutralized separate atom contains just one electrion in equilibrium at its center, it is obvious that, when two atoms overlap so far that the center of one of them is within the spherical boundary of the other, the previous equilibrium of the two electrions is upset, and they must find positions of equilibrium elsewhere than at the centers. Thus, in Fig. 1 given by Kelvin each electrion is at the center of its atom, and is attracted and repelled with equal forces by the neighbouring atom and electrion at its center. In Fig. 2, if E and E' were at the centers

![Fig. 1](image1)

![Fig. 2](image2)
C, C' (C being inside the atom A' which is larger than the atom A), E would be repelled by E' more than it would be attracted by A'. Because both the electrons are supposed to be free, E will move to the right, and on account of its diminished repulsion on E', E' will follow it in the same direction. The equations of equilibrium are easily written down, but not so easily solved.

Calling $e$ the quantity of electricity, positive or negative, in each atom or electron, $\zeta$ the distance between the centers of the atoms, $\alpha$ and $\alpha'$ radii of the two atoms, $x$ and $x'$ the displacements of the electrons from the centers, and $X$ and $X'$ the forces experienced by electrons, we have

$$X = e^2 \left[ -\frac{x}{\alpha^3} + \frac{1}{(\zeta + x - x')^2} - \frac{\zeta + x}{\alpha'^3} \right],$$

$$X' = e^2 \left[ -\frac{x'}{\alpha'^3} + \frac{1}{(\zeta - x')^2} - \frac{1}{(\zeta + x - x')^2} \right].$$

At the equilibrium we have

$$X = X' = 0.$$

These equations are not easily dealt with by frontal attack for the determination of $x$ and $x'$, but may be solved by a method of successive approximation. Let $x_0, x_1, \ldots, x_i$ and $x'_0, x'_1, \ldots, x'_i$ be successive approximations to $x$ and $x'$. Then we have

$$x_{i+1} = \frac{1}{(1/\alpha^3) + (1/\alpha'^3)} \left( \frac{1}{D_i^2} - \frac{\zeta}{\alpha'^3} \right),$$

$$x'_{i+1} = \alpha'^3 \left\{ \frac{1}{(\zeta - x'_i)^2} - \frac{1}{(\zeta + x_{i+1} - x'_i)^2} \right\},$$

where

$$D_i^2 = (\zeta + x_{i+1} - x'_i)^2.$$

Kelvin takes as an example $\alpha = 1, \alpha' = 3$, and makes calculations for the cases of $\zeta = 2.9, 2.8, 2.7, 2.6, 2.5$ finding no convergence for $\zeta = 2.5$. He infers that a...
position of unstable equilibrium is reached for some value of $\zeta$ between 2.7 and 2.6. Namely, if the atoms are brought a little nearer, the equilibrium becomes unstable. Both the electrons jump to the right, $E'$ inside the atom $A$ to settle at a point within it on the left-hand side of its center, and $E$ outside $A'$ to settle at a point still within $A$. If the centers are brought closer and closer together till they coincide, $E$ comes again within $A'$, and the two electrons settle, as shown in Fig. 3, at distances on the two sides of the common center, each equal to

$$\frac{1}{2} \sqrt[3]{\frac{2}{(1/\alpha^3) + (1/\alpha'^3)}}$$

which for the case $\alpha' = 3\alpha$ is 0.622$\alpha$.

Kelvin considers that mutual action of this kind might probably be presented in such diatomic molecules as $O_2$, $N_2$, $H_2$, $CO_2$, $CO$, $SO$, $NaCl$, if each single atom, $O$, $N$, $H$, $Cl$, $C$, $S$, $Na$, had just one electron for its neutralizing number. This may be the case for combinations of two atoms of different quality each of which has one neutralizing electron, as $SO$, for example. How is it, in the case of combination of two atoms of identical quality, as $O_2$, $N_2$, for example? In this case, there must be a strong atomic repulsion preventing stable equilibrium with coincident centers, however strong the attractions by the electrons may be. Because without such a repulsion the two similar atoms would become one, which no possible action in nature could split into two — Kelvin considered so.
Returning to Fig. 3, he discusses the case of pulling the two atoms asunder from the concentric position. It is seen that the electrons will both remain within the smaller atom A, slightly disturbed from equality of distance on the two sides of its center by attractions towards the center of A'. When A' is infinitely distant they will settle at distances each equal to

$$(1/2)\alpha^{3/2} = 0.62996\alpha,$$

on the two sides of the center of A.

If instead of two monoelectronic atoms, two polyelectronic atoms are dealt, after separation the number of electrons in the smaller atom increases and in the larger decreases. Kelvin says that “this is a very remarkable conclusion, pointing to what is probably the true explanation of the first known of the electric properties of matter; attractions and repulsions produced by rubbed amber”. Two ideal solids consisting of assemblages of monoelectronic atoms of largely different sizes would, when pressed and rubbed together and separated, show the properties of oppositely electrified bodies, and the preponderance of the electronic quality would be in the assemblage of which the atoms are smaller. He considers that this is probably the true explanation of the old-known fact that ground glass is negatively charged relatively to polished glass, because the process of polishing might be expected to smooth down the smaller atoms, and to leave the larger atoms more effective in the surface.

Kelvin explains also in this way Erskin Murray’s experimental discovery in 1898 that surfaces of metals become more positive or less negative in the Volta contact electricity scale by being burnished with a steel burnisher.

To illustrate further, he considers two atoms: A', of radius $\alpha'$, the greater, having an electron in it to begin with, and A, of radius $\alpha$, the smaller, void. A is supposed to approach gradually A'. The attraction of A draws the electron from the center of A'. In this case, the force $X$ exerted on the electron E is

$$X = e^2 \left[ \frac{1}{(\zeta - x')^2} = \frac{x'}{\alpha'^3} \right],$$

so that for equilibrium one has

$$\frac{1}{(\zeta - x')^2} = \frac{x'}{\alpha'^3}. \quad (101)$$

This gives an unstable equilibrium. If $\zeta$ is diminished from $2\alpha'$ to $1.89\alpha'$ ($\alpha < 0.89\alpha'$), the electron comes gradually to the distance $0.63\alpha'$ from C'. Its
equilibrium there becomes unstable, and it jumps out of $A'$ toward $A$. "After several oscillations to and fro, perhaps ten or twenty, if it has only quasi inertia due to condensation or rarefaction produced by it in ether, or perhaps many times more, if it has intrinsic inertia of its own, it will settle, with decreasing range of excursions, sensibly to rest within $A$, attracted somewhat from the center by $A'$." If, lastly, $A'$ and $A$ are separated to their original great distance, the electrion will not regain its original position in $A'$, but will come to the center of $A$ and rest there.

In this case more work is gained during the approach than that spent on the separation, and the surplus energy is carried away by ether wave into space. In other words, the system has less potential energy when the electrion is at the center of $A$ than it has when the electrion is at the center of $A'$, as is calculated from the equations given above.

Kelvin considers then static problems of two or more electrions within one atom. If the number of electrions is assumed to be infinite as an extreme case, they can be treated just as Aepinus' electric fluid, and hence they distribute uniformly in the atom, because they take the same distribution as that of the positive electricity. About configuration of a finite and great number of electrions, Kelvin argues by analogy to the case of crystals. Namely, he considers Bravais' homogeneous assemblages of lattice space, and tries finding stable ones. Kelvin takes a configuration which he named equilateral in 1893. It is the one in which each electrion with any two of its twelve next neighbours forms an equilateral triangle. In this case, if all the electrions in the surface layer are left perfectly free, a slight rearrangement among themselves and still slighter among the neighbouring electrions in the interior will bring the multitude to equilibrium. Kelvin says that this subject is of extreme interest, geometrical, dynamical, and physical.

About the case of a moderate number of electrions, he states that for any number of electrions there may be equilibrium with all the electrions on one spherical surface concentric with the atom. To prove this, Kelvin discards for a moment the atom, and imagines the electrions to be attached to ends of equal inextensible strings of which the other ends are fixed to one point $C$. Every string will be stretched in virtue of the mutual repulsions of the electrions, and there will be a configuration or configurations with the electrions on a spherical surface. This is independent of the number of electrions. There is also a configuration of equilibrium in which all the electrions are round one circle in a plane, but it is the sole configuration for two or three electrions.
For four electrons it is seen that they are at the corners of an equilateral tetrahedron.

For five electrons there is a stable equilibrium with three of them in one plane through C, and the other two on both sides of this plane. In another configuration of equilibrium, four of them are in a plane at some distance from C, and the fifth is at the far end of the diameter perpendicular to the plane. Similar configurations may be considered for greater numbers of electrons. Kelvin says that they would be unstable though it is not easy to conclude immediately.

In the case of six electrons they form an octahedron, and in the case of eight a cube. For any even number of electrons there would be two configurations of equilibrium with two halves in two planes at equal distances on the two sides of C, though they are certainly unstable for large numbers. For twelve there comes an equilateral homogeneous assemblage around C, which is Kelvin’s forte, and for twenty a pentagonal dodecahedron.

Returning to the case of an atom, one has electrostatic force towards the center instead of the tension $T$ of the string. In addition radial displacements are possible in this case. Let $e$ be the electric quantity of an electron, and $i$ the number of electrons. One has then $ie^2r/\alpha^3 = T$, where $r$ denotes the radius of the spherical surface on which the electrons lie, and $\alpha$ the radius of the atom. Generally one has $T = q(e^2/r^2)$, where $q$ is a number depending on the number and configuration of the electrons found by geometry. The ratio between two radii becomes hence $r/\alpha = \sqrt[3]{q/i}$. In the case of eight electrons at the corners of a cube, for example, one finds $q = (3/4)(\sqrt{3} + \sqrt{3}/2 + 1/3)$ and therefore $r/\alpha = 0.6756$. With this method Kelvin calculates the work required to remove the electrons to infinite distance for several cases of the number of electrons.

Kelvin thinks that these configurations on one spherical surface are stable possibly even for eight electrons. He says that a pentagonal dodecahedron for twenty electrons does probably not give a stable equilibrium, and for twelve and ten electrons the configurations described in the two paragraphs before the last may probably be unstable. In fact, when the number of electrons exceeds four, one must think of the tendency to be crowded out of one spherical surface, which with very large numbers gives a tendency to uniform distribution throughout the volume of the atom. For example, in the case of five electrons, there is a configuration of equilibrium in which the two electrons lying in one diameter are pushed further from the center than are the three in the equatorial plane. Another configuration, also stable, of five electrons is one at the center.
The proportion of electrons at the center increases with the total number of electrons in an atom.

Explanations of properties of matter — From the theory of atomic structure described above, Kelvin treats the properties of a real body, gaseous, liquid, or solid, constituted by an assemblage of atoms with their electrions. Kelvin states that it follows immediately from his hypothesis, that in a monoatomic gas or in any sufficiently sparse assemblage of single atoms, Faraday’s “conducting power for lines of electric force”, or what is at that time commonly called the specific electro-inductive capacity, or the electro-inductive permeability, exceeds unity by three times the ratio of the sum of the volumes of the atoms to the whole volume of space occupied by the assemblage, whether the atoms be monoelectronic or polyelectronic, and however much the electron, or group of electrions, within each atom is set to vibrate or rotate with each collision, according to the kinetic theory of gases. He takes the case of a single atom in a uniform field of electrostatic force of intensity $F$, $i$ electrons being in the atom. The action of $F$ produces simply displacements of the electrions relatively to the atom, equal and in parallel lines, with therefore no change of force and no rotation. If $x$ denotes the amount of this displacement, the equation for the equilibrium of each electrion is $iex/a^3 = F$ where $a$ is the atomic radius. This gives $iex = a^3F$ for the electric moment of the electrostatic polarization induced in the atom by $F$, and $a^3F$ is equal to the electric moment of the polarization produced in an insulated unelectrified metal globe of radius $a$. This holds whatever the number of electrions in each atom is. Kelvin states that the hypothetical atom hence realizes perfectly for sparse assemblages Faraday’s suggestion (1769) of “small globular conductors” to explain the electro-polarization in solid and liquid insulators.

Denoting by $N$ the number of atoms per unit volume, Kelvin finds $NVa^3F$ as the electric moment of a sparse assemblage of uniform density occupying volume $V$. He says that $N\alpha^3$ is hence the electro-inductive susceptibility suggested by himself. If this is denoted by $\mu$, the electro-inductive permeability $\omega$ is given by

$$\omega = 1 + 4\pi\mu = 1 + 3\left(\frac{N\cdot4\pi\alpha^3}{3}\right),$$

which proves his statement described at the beginning of the last paragraph.

He remarks also that, if vibrating and rotating groups of electrions are included in the demonstration, the influence of $F$ will be the same as if the
assemblage were at rest, because the time average of the position of the center of inertia of the group is taken.

Kelvin states his expectation that the suppositions of forces between electrons and atoms seem to open the way, to a very definite detailed dynamics of electrolysis, of chemical affinity, and of heat of chemical combination. In other words, he thinks that estimates of the actual magnitudes concerned, the number of molecules per cubic centimeters of a gas, the masses of atoms, the diameters of atoms, the electric quantity in an electrion, the mass of an electrion, seem to show the possibility of developing a mechanics of chemistry. However, not being satisfied only with these considerations, he further says — "We might be tempted to assume that all chemical action is electric, and that all varieties of chemical substance are to be explained by the numbers of the electrions required to neutralize an atom or a set of atoms; but we can feel no satisfaction in this idea when we consider the great and wild variety of quality and affinities manifested by the different substances or the different "chemical elements"; and as we are assuming the electrions to be all alike, we must fall back on Father Boscovich, and require him to explain the difference of quality of different chemical substances by different laws of force between the different atoms."

Kelvin treats then solids. He starts from the fact that every solid is, at zero of absolute temperature, a perfect insulator under the influence of electric forces, moderate enough not to pluck electrions out of the atoms. He calls the limiting value of $F$ here indicated for perfect insulation the disruptional force or disruptional intensity. This force is smaller the greater the number of electrions in an atom.

To calculate the electro-inductive permeability of a solid, the static dynamics described above for a gas must be modified, by taking into account forces due to the attractions of neighbouring atoms and the repulsion of their electrions. The induced electric moment will not be in simple proportion to $F$, but will tend to increase more.

When the temperature of a solid is raised to anything under the melting point, the electrions are set to perform wildly irregular vibrations, so that some of them will occasionally be shot out of their atoms. Each electrion thus shot out will quickly either fall back into the atom from which it has been ejected or find its way into another atom. If the body is in an electric field $F$, a considerable proportion of the shot out electrions will find their way in other atoms in the direction in which they are pulled by $F$. The body has now some
degree of electric conductivity, which is greater the higher the temperature. Kelvin explains on this basis T. Gray's experimental result that the electric conductivity of flint glass increases about 800 times between 60°C and 200°C, and the fact that the electric conductivity of rare earths increases enormously above 800°C.

Kelvin explains in the same way the electric conductivity of metals. But in this case he considers the only difference, that the metallic atom is so crowded with electrons that some of them are always being spilled out of each atom by the intermolecular and electronic thermal motions, not only at ordinary temperatures and higher, but even at temperatures less than the absolute temperature 16°. He says 16° because in Dewar's study in 1901 several metals are found to have exceedingly high electric conductivity between the melting and boiling points of hydrogen. He thinks that there is no difficulty in believing that the electrons in each of the metallic atoms are so numerous that, though they rest in stable equilibrium within the atoms at 0° absolute, and may move about within the atom with their irregular thermal motions at 1° absolute, they may between 1° and 2° begin to spill from atom to atom. Thus, he concludes that, like glass and others below 300° absolute, a metal may be an almost perfect electric insulator below 1° absolute. This is of course completely opposite, as we see, to the present-day results and theory of metals.

Problems of crystals — Lastly Kelvin applies his hypothesis to the pyro-electricity and piezo-electricity. He defines a crystal as a homogeneous assemble of bodies in which the distance between centers of nearest neighbours is $10^{-8}$ cm or less. Kelvin explains from his atomic model these phenomena, in which positive and negative electric charges are produced on different parts of the surface of a crystal by change of temperature or pressure. Simplifying the matter, he considers a row of equal tetraelectronic atoms in a straight line. Figure 4 shows a configuration of stable equilibrium of the electrons in this case. The sets of three dots in Fig. 4 indicate trios of electrons at the corners of equilateral triangles, the middle dot in each row being alternately on the far side and the near side of the plane of the paper, which contains the centers of the atoms. Let $C_1, C, C'$ be the centers of the atoms $A_1, A, A'$. The quartet of electrons within $A$ is attracted by $A_1$ with a less force than by $A'$. In other words, $A_1$ attracts all the electrons within $A$ as if it were a positive electric charge $e$ at $C_1$, and similarly in respect to $A'$. There are

\[79\) It is interesting in relation to Kelvin's paper under consideration, that the pyro-electricity was shown as an electric phenomenon by Aepinus in 1756.\]
corresponding smaller differences between the opposite attractions of the more and more remote atoms on the two sides of A. If \( \delta \) is the excess of the sum of the rightwards of these attractions above the leftwards, the geometrical center of the electrons within A is displaced rightwards from C by a distance \( l = a^3 \delta / (16e^2) \).

Imagining a crystal or solid to be built up of parallel rows of atoms such as discussed above, Kelvin thinks that general character of the result will be the same, namely, that throughout the solid, except in a thin superficial layer of perhaps five or ten atoms deep, the whole interior is in stage of homogeneous electric polarization, of which the electric moment per unit volume is \( 4eNl \), where \( N \) is the number of atoms per unit volume. This is, he asserts, the interior molecular condition of a dipolar pyro-electric crystal, which he described in 1860. According to him, the interior electro-polarization becomes necessarily altered, if either by change of temperature or by mechanical stress the distances between the atoms are altered.

Kelvin gives also an explanation of crystals having several dipolar axes.

As we have seen above, Kelvin treated only static cases. In a later paper he supposed potentials of volcanic form in atoms to deal with the problem of \( \beta \)-radioactivity. Kelvin's theory described above gave J. J. Thomson a definite ground to develop his own idea similar to Kelvin's.

§5. Permanence and magnetism of atoms

It is a characteristic of Kelvin's model of the atom without nucleus, that states of static equilibrium are easily found for motionless electrons. Even for the case of one electron static stability is reached. Because in these cases atoms make no radiation, such a model might agree with real situations.

Another characteristic of Kelvin's model of the atom is that the forces acting on electrons are in the direction towards the center of the atom and proportional to their distances from it. This is in agreement with Hertz's
oscillator, which is nothing but a harmonic oscillator that is very simple and ideal in mathematical treatment. This kind of oscillator was used by Planck in his theory of radiation as resonator. Lorentz used also this type of oscillation to explain the Zeeman effect. Because Kelvin gave almost no treatment of the problem of radiation, in his theory no description was given of these things we have mentioned here. Kelvin’s model of atoms without nucleus was, however, possessed of such kind of classical completeness as we have seen above.

When an electron is in motion with acceleration, it radiates electromagnetic energy according to the Maxwell theory of electromagnetism, so that it loses energy. It is not allowed, therefore, to suppose any stable state for an electron moving within an atom. This applies, for example, to the case of an electron making an inner-atomic circular motion. Any atom having one or more electrons in motion will continue to radiate till they cease to move, getting no permanence. Hence it cannot be explained, that some kinds of atoms make no radiation though they have magnetism. At least one electron in an atom must make rotational motion, in order for the atom to have magnetism.

To this problem J. J. Thomson gave an answer in 1903 on the basis of the Maxwell theory of electromagnetism. It was presented in his paper “The Magnetic Properties of Systems of Corpuscles describing Circular Orbits”.

In this paper, he calculates the magnetic field due to a number of negatively electrified corpuscles situated at equal intervals round the circumference of a circle and rotating in one plane with uniform velocity round its center. He does not assume Kelvin’s model of the atom or any central force, but makes his calculations basing only on the premise that the electrons make circular motion. He gets a favorable result that the radiation from the atom decreases rapidly with the increase in the number of electrons, and even for several electrons their rotational motions in the atom can be regarded as stable.

Thomson treats first an electrified particle in rotational motion. Its velocity is assumed to be small compared with the light velocity $V$. It produces a magnetic field whose components $\alpha, \beta, \gamma$ at a distance $r$ from it are given by the equations

$$\alpha = e \left( v \frac{d}{dz} - w \frac{d}{dy} \right) \frac{1}{r},$$

\[81\] Translater's note: In Thomson’s paper the symbols of partial differentiation $\partial/\partial x$, etc. are not used.
\[ \beta = e \left( u \frac{d}{dx} - v \frac{d}{dy} \right) \frac{1}{r}, \quad (103) \]

\[ \gamma = e \left( u \frac{d}{dy} - u \frac{d}{dx} \right) \frac{1}{r}, \]

where \( e \) is the electric charge of the particle and \( u, v, w \) the components of its velocity. These equations cannot be applied to the case of variable motion. When, however, the velocity is periodic, proportional say to \( e^{ipt} \), the equations for the components of the magnetic force can be derived from Eq. (103) by the replacement

\[ \frac{1}{r} \rightarrow e^{-i(p/V)t/r}. \quad (104) \]

This is because the modified equations satisfy differential equations of the type

\[ \frac{d^2 \alpha}{dx^2} + \frac{d^2 \alpha}{dy^2} + \frac{d^2 \alpha}{dz^2} = \frac{1}{V^2} \frac{d^2 \alpha}{dt^2}, \quad (105) \]

and for \( r \) small compared with the wavelength they give the same magnetic force as what would be produced if the particle was moving uniformly with the velocity it possessed at the instant under consideration.

The method described above may be extended to cases more complicated than that of a single particle. If there is a collection of particles separated from each other by distances small compared with the wavelengths of their vibrations, and if in the expression for the magnetic force due to these particles, calculated on the assumption that the velocity is uniform, there is the term

\[ \phi(t) \frac{d^l}{dx^l} \frac{d^m}{dy^m} \frac{d^n}{dz^n} \frac{1}{r}, \quad (106) \]

\( r \) being the distance from a point in the midst of the particle, then if \( \phi(t) \) is proportional to \( e^{ipt} \), the solution, when the variability of the motion is taken into account, will have for the corresponding term

\[ \phi(t) \frac{d^l}{dx^l} \frac{d^m}{dy^m} \frac{d^n}{dz^n} \frac{e^{-(is/V)r}}{r}. \quad (107) \]

Thomson finds first the magnetic field due to a particle with charge \( e \) describing with uniform velocity a circular orbit in the plane \( xy \). Let \( O \) be the center of the orbit, \( P \) the point, not necessarily in the plane \( xy \), at which
the magnetic force is to be determined. If \( a \) is the radius of the orbit, \( \omega \) the angular velocity of rotation, and \( \theta \) the angle \( OQ \) makes with the \( x \) axis, one has

\[
u = -a\omega \sin \theta, \quad v = a\omega \cos \theta, \quad w = 0. \tag{108}\]

When the velocity of the particle is uniform \( \gamma \), the \( z \) component of the magnetic force at \( P \), is given by

\[
\gamma = -eaw \left\{ \sin \theta \frac{d}{dy} + \cos \theta \frac{d}{dx} \right\} \frac{1}{r'}, \tag{109}
\]

where \( r' \) denotes \( PQ \). Denoting \( OP \) by \( r \), one has by expansion

\[
\frac{1}{r'} = \frac{1}{r} - a \left( \cos \theta \frac{d}{dx} + \sin \theta \frac{d}{dy} \right) \frac{1}{r}
+ \frac{a^2}{1 \cdot 2} \left( \cos \theta \frac{d}{dx} + \sin \theta \frac{d}{dy} \right)^2 \frac{1}{r} - \cdots. \tag{110}\]

Hence, writing \( \vartheta \) for

\[
\vartheta = \cos \theta \frac{d}{dx} + \sin \theta \frac{d}{dy}, \tag{111}\]

one has

\[
r = -eaw \left\{ \vartheta \frac{1}{r} - a\vartheta^2 \frac{1}{r} + \frac{a^2}{1 \cdot 2} \vartheta^3 \frac{1}{r} - \cdots \right\}. \tag{112}\]

Thomson passes then to the case in which the particles are in acceleration. Now \( \theta = \omega t \) if the time \( t \) is measured from the instant when the particle is on the \( x \) axis, so that \( \vartheta \) is expressed as follows.

\[
\vartheta = \frac{1}{2} \left\{ e^{i\omega t} \left( \frac{d}{dx} - i \frac{d}{dy} \right) + e^{-i\omega t} \left( \frac{d}{dx} + i \frac{d}{dy} \right) \right\}. \tag{113}\]
Making the transformation of coordinates

$$\xi = x + iy, \quad \eta = x - iy,$$

and taking into account the acceleration of the particle, he shows that the substitution

$$\frac{\gamma n}{r} \to e^{in\omega t} \left( \frac{d}{d\xi} \right)^n e^{-in\omega r/V} r + ne^{i(n-2)\omega t} \left( \frac{d}{d\xi} \right)^{n-1} e^{-i(n-2)\omega r/V} \frac{d}{d\xi} r \left( \frac{d}{d\xi} \right)^2 e^{-i(n-4)\omega r/V} + \ldots$$

into Eq. (112) gives \(\gamma\). Of \(\alpha, \beta\) he makes similar calculations.

The resultant expressions consist of two parts: (1) periodic terms proportional to some power of \(e^{i\omega t}\), and (2) terms independent of time. In considering the magnetic force due to the particles, Thomson takes these terms separately and retains in each only the lowest power of \(a\).

He begins with the case of a single particle and takes first the terms independent of time. These terms are given by the following equations.

$$\begin{align*}
\gamma &= \frac{1}{2} ea^2 \omega \left( \frac{d^2}{dx^2} + \frac{d^2}{dy^2} \right) \frac{1}{r} = -\frac{1}{2} ea^2 \omega \frac{d^2}{dz^2} \frac{1}{r}, \\
\alpha &= -\frac{1}{2} ea^2 \omega \frac{d}{dx} \frac{d}{dz} \frac{1}{r}, \\
\beta &= -\frac{1}{2} ea^2 \omega \frac{d}{dy} \frac{d}{dz} \frac{1}{r},
\end{align*}$$

that is, the part of the magnetic force which is independent of the time is that due to a small magnet of moment \(ea^2\omega/2\) magnetized in the direction of the \(z\) axis.

In the lowest order the periodic terms are

$$\begin{align*}
\gamma &= ea \omega \left\{ \frac{x}{r^3} \cos \left( \omega t - \frac{\omega r}{V} \right) + \frac{y}{r^3} \sin \left( \omega t - \frac{\omega r}{V} \right) \right. \\
&\quad \left. - \frac{xw}{Vr^2} \sin \left( \omega t - \frac{\omega r}{V} \right) + \frac{yw}{Vr^2} \cos \left( \omega t - \frac{\omega r}{V} \right) \right\}, \\
\alpha &= ea \omega \left\{ -\frac{z}{r^3} \cos \left( \omega t - \frac{\omega r}{V} \right) + \frac{zw}{Vr^2} \sin \left( \omega t - \frac{\omega r}{V} \right) \right\}, \\
\beta &= ea \omega \left\{ -\frac{z}{r^3} \sin \left( \omega t - \frac{\omega r}{V} \right) - \frac{zw}{Vr^2} \cos \left( \omega t - \frac{\omega r}{V} \right) \right\}.
\end{align*}$$
When $r \gg V/\omega$, the terms involving the lowest powers of $1/r$ are the most important. Thus, if $Oz$ is the normal drown from $O$ to the plane of the orbit, the component of the magnetic force at $P$ in the plane $POz$ at right-angles to $OP$ and that at right-angles to the plane $POz$ are, respectively,

\[
\begin{align*}
- \frac{e\omega^2}{Vr} \sin \left( \omega t - \frac{\omega r}{V} - \phi \right), \\
\frac{e\omega^2}{Vr} \cos \theta \cos \left( \omega t - \frac{\omega r}{V} - \phi \right),
\end{align*}
\]

(118)

where $\theta$ is the angle $OP$ makes with $Oz$, and $\phi$ the angle the plane $POz$ makes with the plane of $xz$.

As Thomson points out, one may say in the language of the electromagnetic theory of light that the rotating particle produces a wave of elliptically polarized light, the ratio of the axes of the ellipse being $\cos \theta$, and accordingly along the normal to the orbit one has circularly polarized light, while in the plane of the orbit the light is plane polarized.

Along with these magnetic forces there is an electric force, of which the component in the plane $POz$ at right angles to $OP$ and that at right-angles to the plane $POz$ are, respectively,

\[
\begin{align*}
\frac{e\omega^2}{r} \cos \theta \cos \left( \omega t - \frac{\omega r}{V} - \phi \right), \\
\frac{e\omega^2}{Vr} \sin \left( \omega t - \frac{\omega r}{V} - \phi \right).
\end{align*}
\]

(119)

Thomson calculates then by means of the Poynting vector the rate at which energy is streaming through unit area at $P$, getting

\[
\frac{1}{4\pi} \frac{e^2 a^2 \omega^4}{Vr^2} \left\{ \cos^2 \theta \cos^2 \left( \omega t - \frac{\omega r}{V} - \phi \right) + \sin^2 \left( \omega t - \frac{\omega r}{V} - \phi \right) \right\}.
\]

The mean value of this is

\[
\frac{1}{8\pi} \frac{e^2 a^2 \omega^4}{Vr^2} (1 + \cos^2 \theta).
\]

Integrating this over the sphere through $P$, Thomson finds the mean rate at which the rotating corpuscle is emitting energy to be

\[
\frac{2}{3V} e^2 a^2 \omega^4 = \frac{2e^2}{3V} \text{(acceleration of the particle)}^2.
\]

(120)
Passing to the case of $p$ particles separated by equal angular intervals rotating with uniform velocity $\omega$ round a circle, Thomson supposes that the particle (1) makes at time $t$ an angle $\omega t$ with the $x$ axis, the particle (2) an angle $\omega t + 2\pi/p$, the particle (3) an angle $\omega t + 2(2\pi/p)$, and so on. Hence, if $\gamma_1, \gamma_2, \gamma_3 \cdots$ are the magnetic forces parallel to the $z$ axis due to the particle (1), (2), (3) \cdots respectively, $\gamma_1$ is given by the expression already found, $\gamma_2$ is got from by writing $\omega t + 2\pi/p$ for $\omega t$, $\gamma_3$ by writing $\omega t + 2(2\pi/p)$ for $\omega t$, and so on. The total magnetic force parallel to the $z$ axis due to the $p$ particles is $\gamma_1 + \gamma_2 + \gamma_3 + \cdots + \gamma_p$. The term independent of time is same in $\gamma_1, \gamma_2, \gamma_3 \cdots$, and hence for the $p$ particles it is

$$-pea^2 \omega \frac{d^2}{dz^2} \frac{1}{r}.$$  

Of the periodic terms, the one corresponding to $e^{i\omega t}$ in $\gamma_1$ is in $\gamma_2 e^{i(\omega t+2\pi/p)}$, in $\gamma_3 e^{i(\omega t+2\cdot2\pi/p)}$, and so on, and hence the term corresponding to $\gamma_1 + \gamma_2 + \gamma_3 + \cdots$ is

$$e^{i\omega t} + e^{i(\omega t+2\pi/p)} + e^{i(\omega t+2\cdot2\pi/p)} + e^{i(\omega t+3\cdot2\pi/p)} + \ldots.$$  

The expression vanishes unless $n$ is a multiple of $p$, when it equals $pe^{i\omega t}$.

Hence the largest periodic term in the expression for the magnetic force due to the $p$ corpuscles is that corresponding to the term $e^{ip\omega t}$ in $\gamma_1$, and its magnitude is $p$ times this term, that is,

$$(-1)^p \frac{\omega a^p}{1 \cdot 2 \cdot 3 \cdots (p-1)} \times p \left\{ e^{ip\omega t} \left( \frac{d}{d\xi} \right)^p \frac{e^{ip\omega r/V}}{r} + e^{-ip\omega t} \left( \frac{d}{d\eta} \right)^p \frac{e^{ip\omega r/V}}{r} \right\}$$

$$= (-1)^p \frac{\omega a^p p}{1 \cdot 2 \cdot 3 \cdots (p-1)} \cdot \frac{1}{2^p} \left\{ (x - iy)^p \left( \frac{d}{rdr} \right)^p \frac{e^{i(p\omega t - p\omega r/V)}}{r} + (x + iy)^p \left( \frac{d}{rdr} \right)^p \frac{e^{-i(p\omega t - p\omega r/V)}}{r} \right\}.$$  

The $x$ and $y$ components of the force due to $p$ particles are obtained in similar ways.

At a distance large compared with the wavelength the terms proportional to $1/r$ are most important, the other terms becoming insignificant. Hence, at a point $P$ the component of the magnetic force in the plane $POz$ at right-angles to $OP$ is
The components of the electric force are \( VL \) in the plane \( POz \) and \( VM \) at right-angles to it.

These expressions represent also elliptically polarized light, the ratio of the axes being, as in the case of one particle, \( \cos \theta \). In the present case, however, since \( L = M = 0 \) for \( \theta = 0 \), the intensity of the light vanishes along the normal \( Oz \) to the plane of the orbit. Thomson says that "thus a system of two or more particles rotating uniformly in a circular orbit does not give out any light along the axis of that orbit, a point which has to be kept in mind when considering the interpretation of the Zeeman effect". He points out that as \( p \) increases the vibrations tend to become confined to the plane of the orbit.

The rate at which energy is streaming through unit area of the surface at \( P \) is by Poynting's theorem equal to

\[
\frac{V}{4\pi} \left( L^2 + M^2 \right).
\]

The average rate at which energy is streaming through the surface of a sphere with its center at \( O \) becomes therefore

\[
\frac{V}{4} \left( \frac{2\epsilon \omega p}{1 \cdot 2 \cdot 3 \cdots (p - 1)} \right)^2 \left( \frac{\alpha \omega}{2V} \right)^2 \int_0^\pi \sin^{2p-2} \theta (1 + \cos^2 \theta) \sin \theta d\theta
\]

\[
= \frac{V}{4} \left( \frac{2\epsilon \omega p}{1 \cdot 2 \cdot 3 \cdots (p - 1)} \right)^2 \left( \frac{\alpha \omega}{2V} \right)^2 \frac{2 \cdot 4 \cdots (2p - 2) \cdot (2p + 2)}{1 \cdot 3 \cdots (2p + 1)}.
\]  

From this expression it is seen that, when the velocity \( \alpha \omega \) of the particle is small compared with the light velocity \( V \), the rate of energy radiation diminishes very rapidly as the number of particles increases. Thomson brings out
this by the following table, which gives the average radiation per particle in terms of the radiation from a single particle, for two cases of the velocity of the particle.

<table>
<thead>
<tr>
<th>Number of particles</th>
<th>( aw = V/10 )</th>
<th>( aw = V/10^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>( 9.6 \times 10^{-2} )</td>
<td>( 9.6 \times 10^{-4} )</td>
</tr>
<tr>
<td>3</td>
<td>( 4.6 \times 10^{-3} )</td>
<td>( 4.6 \times 10^{-7} )</td>
</tr>
<tr>
<td>4</td>
<td>( 1.7 \times 10^{-4} )</td>
<td>( 1.7 \times 10^{-10} )</td>
</tr>
<tr>
<td>5</td>
<td>( 5.6 \times 10^{-5} )</td>
<td>( 5.6 \times 10^{-13} )</td>
</tr>
<tr>
<td>6</td>
<td>( 1.6 \times 10^{-7} )</td>
<td>( 1.6 \times 10^{-17} )</td>
</tr>
</tbody>
</table>

From this table one sees that the radiation even from a group of 6 particles is considerably weak.

Thomson treats next the magnetism of matter, and studies the action of an external magnetic field on a ring of rotating particles. He says that though “it might seem at first sight as if a body whose atoms contained systems of such rotating particles ought to be strongly magnetic, the following investigation of the action of an external field on such a system shows that this is not the case”. In other words, Thomson calculates the motion of an atomic electron, which moves under an arbitrary central force \((X, Y, Z)\) and under the influence of an external magnetic field \(H\), and inserts it into the expression for \(\alpha\). Taking the lowest order approximation and taking the average over the planes of the orbit as well as the positions of the particle, he finds that the resultant is zero. Hence, Thomson concludes that “we cannot explain the magnetic or diamagnetic properties of bodies by the supposition that the atoms consist of charged particles describing closed periodic orbit under the action of a force proportional to the distance from a fixed point”, and further that “the same absence of diamagnetic or magnetic properties will persist whatever be the law of force, provided the force is central and there is no dissipation of energy”.

When there is dissipation of energy, Thomson represents it by supposing that the particles are resisted by a force \((kdx/dt, kdy/dt, kdz/dt)\) proportional to the velocity. For a uniform magnetic field \(H\) in the direction of the \(x\) axis and a central force \((X, Y, Z)\), the equations of motion become
The Formation of Atomic Models

\[
\begin{align*}
    m \frac{d^2 x}{dt^2} &= X - k \frac{dx}{dt}, \\
    m \frac{d^2 y}{dt^2} &= Y - k \frac{dy}{dt} - He \frac{dz}{dt}, \\
    m \frac{d^2 z}{dt^2} &= Z - k \frac{dz}{dt} + He \frac{dy}{dt}.
\end{align*}
\]

(125)

Inserting the solution of these equations into the expression for \( \alpha \), and retaining terms in the lowest order approximation, Thomson shows that the contribution of the particle to the coefficient of magnetization is the mean value of the expression

\[
-\frac{1}{4} \frac{e^2}{m} e^{-\frac{(k/m)t}{4}} \int_0^t e^{\frac{(k/m)t}{4}} \frac{d}{dt}(y^2 + z^2)dt,
\]

which is the coefficient of the term proportional to \( H \). If the motion of the particles is strongly retarded the particle will tend to fall in to the center of the orbit. In this case \( d(y^2 + z^2)/dt \) is therefore negative, and the mean value of the above expression will be positive. Hence, the contribution of this particle to the coefficient of magnetization will be positive, and a collection of such particles will be paramagnetic.

Because in a paramagnetic system there would thus be a continual transference of energy from the atoms to the surrounding, Thomson says that “this would tend to raise the temperature of the system”. He then mentions — “I am not aware that any experiments have been made to find whether the temperature in the middle of a mass of a magnetic substance like iron whose surface is kept at a constant temperature differs from the temperature inside a mass of a non-magnetic substance like brass whose surface is kept at the same temperature. I hope, however, soon to be able to test this point.”

Thomson’s theory of magnetic substances based on his theory of atoms did not give right results, and the first step to the theory of magnetic substances was given in 1905 by Langevin\(^{82}\) from the standpoint of atomism, who calculated correctly the Larmor precession by using also the classical theory. The correct theory inclusive of that of ferromagnetic substances was of course made clear first by means of quantum mechanics. Thomson’s theory should anyway be noted as the first theory of magnetic substances based on atomism.

\(^{82}\)P. Langevin, J. de Phys. 4 (1905), 678; Ann. de Chem. et Phys. 5 (1905), 70.
Thomson’s study of the magnetism of atoms we have described above was grounded on W. Voigt’s study\(^{83}\) published one year before Thomson’s. Voigt started from a model of the interior of atoms, which was similar to that in Lorentz’s theory of the Zeeman effect, but did not succeed in deriving the magnetism of atoms. It was P. Langevin\(^ {84}\) who succeeded first in that, as we have mentioned above. Langevin was able to relate the Larmor precession clearly to diamagnetism. This is a noticeable result in the classical theory.

Langevin says — “Those who have tried to pursue this idea have till now found it sterile (Thomson and Voigt). On the contrary, I have found that it is possible differently from their opinion to give by the hypothesis of electrons an explanation for para- and dia-magnetism, and to get the results which are in accord with various laws of magnetism established by M. P. Curie (1895).” In the following we shall give an outline of Langevin’s assertion.

When a particle with the charge \( e \) has a velocity \( v \), it is equivalent to an element of the moment of current \( ev \). The magnetic field at great distances from an electron is, being calculated from the electric current which it produces when it describes an orbit of area \( S \) with a period \( \tau \), the same as if there were a magnet with the magnetic moment

\[
M = \frac{eS}{\tau}
\]

in the direction normal to the plane of the orbit. We have already mentioned this above in relation to Eq. (116), from which we see that \( M = (1/2)ea^2\omega \).

Langevin asserts that the particular current corresponding to each electron in a molecule and the magnetic moment due to it, disappear or remain in accordance with the symmetry of the molecular structure.

When an aggregate of molecules of such a kind is acted on by an external magnetic field, the current due to each particle is subjected to a deformation in a direction independent of the action. The direction of such a deformation corresponds always to diamagnetism. In the case of the circular orbit, the increase in magnetic moment is, if \( H \) is the component of the external magnetic field normal to the plane of the orbit,

\[
-\Delta M = -\left( \frac{He^2}{4\pi m} \right) S,
\]

\( m \) being the mass of an electron.


\(^{84}\)Langevin’s study was published first in Compt. Rend. (1904), 1206.
The Thomson model of the atom

As is mentioned in §4, J. J. Thomson, who had already in 1897 an idea similar to that in Kelvin’s model of the atom, advanced the model. Thomson developed his model in his lecture “Electricity and Matter” given at the Yale University in May, 1903.

In this lecture, Thomson stresses that, because the mass of a corpuscle is lighter than that of any atom, corpuscles are the constituents of every atom of various kinds, that is, atoms of various kinds of matter have something in common and every atom of any chemical element is composed of simpler systems. Thomson mentions that such a way of thinking came in various forms to former chemists. He refers to W. Prout, an English physicist (1785–1850), who urged in 1815 that every atom of various kinds of elements was composed of hydrogen atoms. Thomson says that “if this were so the combining weight of all the elements would, on the assumption that there was no loss of weight when the atoms of hydrogen combined to form the atom of some other element, be integers; a result not in accordance with observation”.

In order to avoid this disagreement, Dumas (1800–84, French chemist) supposed that the fundamental atom should not be the hydrogen atom, but be a smaller atom having a mass of a half or third of that of a hydrogen atom. He considered that the periodic law of Newlands and Mendeleeff indicated this supposition. He pointed out also that the combination of the atomic weights of some elements gave the atomic weight of an element of a kind similar to

---

them, as in the case, for example, of Na whose atomic weight 23 is given by the arithmetic mean of those of Li and K, 7 and 39 respectively. On this fact he supposed that atoms of various elements should have something common to them.

Thomson considers that this is indicated also by the fact of spectrum, that the spectral structures of the elements of the same family in the periodic table are similar to each other, and says that “indeed spectroscopic evidence alone has led Sir Norman Lockyer for a long time to advocate the view that the elements are really compounds which can be dissociated when the circumstances are suitable”. This point is just as we have described in the section about line spectra. Thomson relates this point further to the radioactivity, saying that “there seems good reasons for believing that radio-activity is due to changes going on within the atoms of the radio-active substances”.

About his model of the atom Thomson gives himself the following valuation — “It may thus not be superfluous to consider the bearing of the existence of corpuscles on the problem of the constitution of the atom; and although the model of the atom to which we are led by these considerations is very crude and imperfect, it may perhaps be of service by suggesting lines of investigations likely to furnish us with further information about the constitution of the atom”.

Thomson proposes the question: of what character every simpler system would be if an atom is supposed to be an aggregate of simpler systems. It is natural to think the corpuscle to be the constituent of a fundamental system, since the mass of the corpuscle is much smaller than those of atoms. The corpuscle has also a definite charge of negative electricity. Hence Thomson asserts that “since with any charge of electricity we always associate an equal charge of the opposite kind, we should expect the negative charge on the corpuscle to be associated with an equal charge of positive electricity”. He supposes thus as the fundamental system an electric doublet which consists of a negative corpuscle at one end and a positive electric charge of equal quantity at the other end, both ends being connected by lines of electric force. He thinks that the lines of electric force possesses a material existence, that is, an electromagnetic mass. Thomson considers that the positive electric charge extends for a volume much larger than that of the corpuscle, as Kelvin did. In the fundamental system therefore the lines of electric force gather near the corpuscle much more densely than at the other parts. According to Thomson, “the quantity of ether bound by the lines of force, the mass of which we regard as the mass of the system, will be very much greater near the corpuscle than elsewhere”. He
thinks thus that the mass of the system is entirely electromagnetic one. Now, since the size of the corpuscle is much smaller than that of the positive electric charge, it results that the mass of the system is due to the mass of the ether combined near the corpuscle, and therefore does not depend on the position of the positive electric charge. Thomson calculated already the electromagnetic mass to find the value $2e^2/3a$ for it, $a$ being the radius of the corpuscle. He gets thereby the value of $a$ as

$$a \sim 10^{-13} \text{ cm}$$

from the knowledge of cathode rays.

Such electric doublets, like small magnets, attract each other to combine. Thomson thinks that this combination results in, "like the Aepinus atom of Lord Kelvin", a sphere charged uniformly and positively, inside which much smaller corpuscles charged negatively move, under the action of an electric central force proportional to the distance of each corpuscle from the center. The number of the corpuscles in a sphere is that of the fundamental systems making up this combination, and the total amount of negative charge of the corpuscles is equal to the amount of positive charge of the sphere.

In this way, Thomson considers that the mass of an atom is mainly due to the electrons in it. Every atom is therefore considered an aggregate of a large number of electrons. He got this idea already in 1897, as we have mentioned before. In his book under consideration ("Electricity and Matter", p. 110) he gives the relation

$$nm = M,$$

for the number $n$ of the electrons in an atom and the masses $m$ and $M$ of the corpuscle and the atom respectively. He supposes $n = 1000$ for the hydrogen atom. About this idea, Thomson mentions also in his paper in the next year\textsuperscript{86} as follows — "We suppose that the mass of an atom is the sum of the masses of the corpuscles it contains, so that the atomic weight of an element is measured by the number of the corpuscles in the atom."

Thomson leaves this idea later in 1907,\textsuperscript{87} because the study of the number of atomic electrons and that of positive rays, both being made by him, force him to accept some mass in addition to the electromagnetic mass, on one hand,

\begin{footnotesize}
\textsuperscript{86} J. J. Thomson, Phil. Mag. 7 (1904), 237.
\textsuperscript{87} J. J. Thomson, "Corpuscular Theory of Matter" (1907), pp. 27 and 162.
\end{footnotesize}
and let him know that the number of atomic electrons is not very large but nearly equal to the atomic weight, on the other hand. We shall discuss this point later.

About the model given in his paper\(^{86}\) Thomson calculates the condition for its stability and the modes of vibration. In the year before Nagaoka makes a similar calculation, which is published\(^{88}\) in Phil. Mag. after Thomson’s paper under consideration. Both calculations follow the example of Maxwell’s famous calculation\(^{89}\) of the condition for the stability of the rings of Saturn.

Each of these calculations aims to get the explanation of spectral laws, as one of its ends, by knowing the modes of vibration. This is quite the same problem as that in Rayleigh’s study we have described already in Sec. 1, §3 of the present chapter, from which one can see that in this way it is impossible to obtain spectral laws.

Now, the title of Thomson’s paper under consideration is “On the Structure of the Atom: an Investigation of the Stability and Periods of Oscillation of a number of Corpuscles arranged at equal intervals around the Circumference of a Circle; with Application of the results to the Theory of Atomic Structure”.

He supposes that in equilibrium \(n\) particles, each carrying a negative electric charge \(e\), are arranged at equal angular intervals round the circumference of a circle of radius \(a\), inside a uniformly electrified sphere. Let the radius of the sphere be \(b\), and the total positive electric charge in the sphere \(\nu e\). The radial attraction on a particle due to the positive charge is then \(\nu e^2 a/b^2\). If the particles are at rest this attraction must be balanced by the repulsion exerted by the other particles. The repulsion along \(OA\) exerted on a particle at \(A\), \(O\) being the center of the radius, by one at \(B\) is, if \(OA = OB\), equal to

\[
\frac{e^2}{AB^2} \cos OAB = \frac{e^2}{4OA^2 \sin(AOB/2)}.
\]

When \(n\) particles are arranged at equal angular intervals \(2\pi/n\) round the circumference, the radial repulsion on one particle due to the other \((n - 1)\) particles is

\[
\frac{e^2}{4a^2} \left\{ \csc \frac{\pi}{n} + \csc \frac{2\pi}{n} + \csc \frac{3\pi}{n} + \cdots + \csc \frac{(n - 1)\pi}{n} \right\}.
\]

\(^{86}H.\ Nagaoka,\ Phil.\ Mag.\ 7\ (1904),\ 445.\)
\(^{89}J.\ C.\ Maxwell,\ “The\ Collected\ Works”,\ Vol.\ I.\)
Let the quantity in the braces be denoted by $S_n$. Because this radial repulsion is to be balanced by the radial attraction, one has

$$\frac{\nu e^2 a}{b^3} = \frac{e^2}{4a^2} S_n,$$

or

$$\frac{a^3}{b^3} = \frac{S_n}{4\nu}.$$  \hspace{1cm} (126)

In the case of $\nu = n$, $a/b$ takes the following values:

\[
\begin{array}{cccccc}
 n & 2 & 3 & 4 & 5 & 6 \\
 a/b & 0.5 & 0.5773 & 0.6208 & 0.6505 & 0.6726 \\
\end{array}
\]

If the ring of particles (corpuscles) is rotating with an angular velocity $\omega$, the condition for steady motion is

$$\frac{\nu e^2 a}{b^3} = m\omega^2 + \frac{e^2}{4a^2} S_n,$$

or

$$\frac{\nu a^3}{b^3} = \frac{ma^2 \omega^2}{e^2} + \frac{S_n}{4},$$

$m$ being the mass of a corpuscle.

Thomson proceeds then to find the forces acting on a corpuscle (electron) when the corpuscles are slightly displaced from their positions of equilibrium. Let the position of the electrons be fixed by the polar coordinates $r$, $\theta$ in the plane of the undisturbed orbit, and by the displacement $z$ at right angles to this plane. Let $r_s$, $\theta_s$, $z_s$ be the coordinates of the $s$th electron. When the electrons are slightly displaced from their positions of equilibrium, $\rho_s$ defined by $r_s = a + \rho_s$ is small compared with $a$. Also $z_s$ is small compared with $a$, and $\phi_s$'s given by $\theta_s - \theta_{s-1} = 2\pi/n + \phi_s - \phi_{s-1}$ are small quantities.

The radial repulsion exerted by the $s$th electron on the $p$th is

$$-e^2 \frac{d}{dr_p} \left\{ \frac{1}{r_p^2 + r_s^2 - 2r_pr_s \cos(\theta_p - \theta_s) + (z_p - z_s)^2} \right\}^{1/2}.$$
Expanding this, retaining only the first powers of $\rho$, $\phi$ and $z$, he finds that if $R_{ps}$ is the repulsion it is given by

$$R_{ps} = \frac{e^2}{4a^2 \sin \psi} \left\{ 1 - \frac{\rho_p}{a} \left( \frac{3}{2} - \frac{1}{2 \sin^2 \psi} \right) - \frac{\rho_s}{2} \left( \frac{1}{2} + \frac{1}{2 \sin^2 \psi} \right) - \frac{1}{2} (\phi_s - \phi_p) \cot \psi \right\},$$

where $\psi = (p - s)\pi/n$.

The tangential force $\Theta_{ps}$ tending to increase $\Theta_p$ is

$$\Theta_{ps} = \frac{e^2}{4a^2 \sin^2 \psi} \left\{ 1 - \frac{3\rho_s}{2a} - \frac{\rho_s}{2a} - (\phi_s - \phi_p) \left( \cot \psi + \frac{1}{2} \tan \psi \right) \right\}.$$

The expansion of this of the same kind as the above one gives

$$\Theta_{ps} = \frac{e^2 \cos \psi}{4a^2 \sin^2 \psi} \left\{ 1 - \frac{3\rho_s}{2a} - \frac{\rho_s}{2a} - (\phi_s - \phi_p) \left( \cot \psi + \frac{1}{2} \tan \psi \right) \right\}.$$

The force $Z_{ps}$ at right angles to the undisturbed plane of the orbit is

$$Z_{ps} = \frac{e^2}{8a^2 \sin \psi} (z_p - z_s).$$

The total radial force $R_p$ exerted on the $p$th electron takes then the form

$$R_p = \sum R_{ps} = \frac{e^2}{4a^2} S - \rho_p A' - \sum \rho_{p+s} A_s - a \sum \phi_{p+s} B_s,$$

where $S$, $A'$ are some functions of $\sin(\pi/n)$, and $A_s$, $B_s$ some ones of $\sin(s\pi/n)$.

The tangential force $\Theta_p$ acting on the $p$th electron is

$$\Theta_p = \sum \Theta_{ps} = \sum \rho_{p+s} B_s - a \phi_p C + a \sum \phi_{p+s} C_s,$$

while $Z_p$, the force at right angles to the plane, is

$$Z_p = \sum Z_{ps} = z_p D - \sum z_{p+s} D_s,$$

where $C$, $D$ and $C_s$, $D_s$ are functions similar to $A'$ and $A_s$, respectively.
The equations of motion of the $p$th particle are thus obtained in the form

$$m \left( \frac{d^2 r_p}{dt^2} - r_p \left( \frac{d \theta_p}{dt} \right)^2 \right) = -\frac{\nu e^2 r_p}{b^3} + R_p,$$

(127a)

$$m \left( r_p \frac{d^2 \theta_p}{dt^2} + 2 \frac{dr_p}{dt} \frac{d \theta_p}{dt} \right) = \Theta_p,$$

(127b)

$$m \frac{d^2 z_p}{dt^2} = -\frac{\nu e^2}{b^3} z_p + Z_p.$$  

(127c)

Retaining only the first powers of small quantities, one gets from these equations, if $\omega$ is the steady value of $d\theta/dt$,

$$\frac{\nu e^2 a}{b^3} = m\omega^2 + \frac{e^2}{4a^2} S.$$

From Eq. (127a) one then has

$$m \frac{d^2 \rho_p}{dt^2} - 2m\omega \frac{d \theta_p}{dt} = \left( m\omega^2 - \frac{\nu e^2}{b^3} \right) + R_p - \frac{e^2}{4a^2} S.$$

If $\rho_p$ and $\theta_p$ are periodic, that is, function of time of the form $e^{i\omega t}$, this equation may be written

$$(A - mq^2) \rho_p + A_1 \rho_{p+1} + A_2 \rho_{p+2} + \cdots$$

$$- 2m\omega iq \phi_p + aB_1 \phi_{p+1} + aB_2 \phi_{p+2} + \cdots = 0,$$

(128a)

where $A = (e^2/4a^3)S + A'$, $(p = 1, 2, 3, \cdots n)$. In a similar way Eq. (127b) gives

$$2m\omega iq \frac{\rho_p}{a} - B_1 \frac{\rho_{p+1}}{a} - B_2 \frac{\rho_{p+2}}{a}$$

$$+ \cdots (C - mq^2) \phi_p - C_1 \phi_{p+1} - C_2 \phi_{p+2} - \cdots = 0,$$

(128b)

$(p = 1, 2, 3, \cdots n)$.

In order to solve Eqs. (128a) and (128b), let $\tilde{\omega}$ be any root of the equation $x^n = 1$. Eq. (128a) is then satisfied by

$$\rho_2 = \tilde{\omega} \rho_1, \rho_3 = \tilde{\omega} \rho_2, \rho_4 = \tilde{\omega} \rho_3, \phi_2 = \tilde{\omega} \phi_1, \phi_3 = \tilde{\omega} \phi_2, \phi_4 = \tilde{\omega} \phi_3 \cdots$$
provided that
\[
\rho_1(A - mq^2 + \bar{\omega}A_1 + \bar{\omega}^2A_2 + \cdots + \bar{\omega}^{n-1}A_{n-1}) \\
+ \phi_1a(-2im\omega q + \bar{\omega}B_1 + \bar{\omega}^2B_2 + \cdots + \bar{\omega}^{n-1}B_{n-1}) = 0, \tag{129a}
\]
while Eq. (128b) is satisfied by the same values provided
\[
\rho_1(2im\omega q - \bar{\omega}B_1 - \bar{\omega}^2B_2 - \cdots - \bar{\omega}^{n-1}B_{n-1}) \\
+ \phi_1a(C - mq^2 - \bar{\omega}C_1 - \bar{\omega}^2C_2 - \cdots - \bar{\omega}^{n-1}C_{n-1}) = 0. \tag{129b}
\]
For both Eqs. (129a) and (129b) to be satisfied simultaneously, it results from the elimination of \(\rho_1\) and \(\phi_1\) that
\[
((A - mq^2) + \bar{\omega}A_1 + \bar{\omega}^2A_2 + \cdots + \bar{\omega}^{n-1}A_{n-1}) \\
\times (C - mq^2 - \bar{\omega}C_1 - \bar{\omega}^2C_2 - \cdots - \bar{\omega}^{n-1}C_{n-1}) \\
= -(2im\omega q - \bar{\omega}B_1 - \bar{\omega}^2B_2 - \cdots - \bar{\omega}^{n-1}B_{n-1}). \tag{130}
\]
Let \(\omega\) be expressed as
\[
\omega = \cos \frac{2k\pi}{n} + i\sin \frac{2k\pi}{n} \quad (k = 0, 1, \ldots, n - 1),
\]
and \(L_k, N_k, M_k\) defined as
\[
L_k = \bar{\omega}A_1 + \bar{\omega}^2A_2 + \cdots + \bar{\omega}^{n-1}A_{n-1},
\]
\[
N_k = \bar{\omega}C_1 + \bar{\omega}^2C_2 + \cdots + \bar{\omega}^{n-1}C_{n-1},
\]
\[
iM_k = \bar{\omega}B_1 + \bar{\omega}^2B_2 + \cdots + \bar{\omega}^{n-1}B_{n-1}.
\]
Because \(C\) is equal to the value of \(N_k\) for \(k = 0\), let \(C\) be written as \(C = N_0\). \(A\) is also expressed as \(A = (3e^2/4a^3)S - L_0\). Introducing these expressions into Eq. (130), one gets
\[
\left(\frac{3e^2}{4a^3}S + L_k - L_0 - mq^2\right)(N_0 - L_k - mq^2) = (M_k - 2m\omega q)^2. \tag{131}
\]
k in this equation may have any value from 0 to \((n - 1)\). If \((n - k)\) is written for \(k\) in this equation, the values of \(q\) differ only in sign, and so give the same frequencies. Thus all the values of \(q\) can be obtained by \(k = 1, 2, \cdots (n - 1)/2\) if \(n\) is odd, or \(k = 1, 2, \cdots n/2\) if \(n\) is even. If \(n\) is odd, there are \((n + 1)/2\) equations of the type of Eq. (131). When \(k = 0\), \(M_k = 0\) and Eq. (131)
The Formation of Atomic Models

reduces to a quadratic equation, so that the number of roots of these \((n+1)/2\) equations is \(4 \times (n+1)/2 - 2 = 2n\). If \(n\) is even, there are \((n/2 + 1)\) equations, but as \(M_k = 0\) for \(k = 0\) and \(k = n/2\), two of them reduce to quadratics. The number of roots of these equations is therefore \(4(n/2 + 1) - 4 = 2n\). Thus in each case the number of roots is equal to \(2n\), the number of degrees of freedom of the particles in the plane of the undisturbed orbit.

Thomson considers now the motion at right angles to this plane. From Eq. (127c) the equation

\[
m \frac{d^2 z_p}{dt^2} = -\frac{\nu e^2}{b^3} z_p + D z_p - \sum D_s z_{p+s}
\]

is obtained. On the supposition \(z_p \sim e^{iqt}\), this reduces to

\[
z_p \left(\frac{\nu e^2}{b^3} - D - mq^2\right) + \sum D_s z_{p+s} = 0. \quad (128c)
\]

The solution of Eq. (128c) is

\[z_2 = \tilde{\omega} z_1, \quad z_3 = \tilde{\omega} z_2, \quad z_4 = \tilde{\omega} z_3, \ldots\]

and

\[
\frac{\nu e^2}{b^3} - D - mq^2 + \tilde{\omega} D_1 + \tilde{\omega}^2 D_1 + \cdots + \tilde{\omega}^{n-1} D_{n-1} = 0. \quad (132)
\]

Putting in a similar way to the above cases

\[P_k = \tilde{\omega} D_1 + \tilde{\omega}^2 D_1 + \cdots + \tilde{\omega}^{n-1} D_{n-1}\]

and noticing that \(D = P_0\), one has from Eq. (132)

\[
\frac{\nu e^2}{b^3} + P_k - P_0 - mq^2 = 0.
\]

Putting \(k = 0, 1, \cdots n-1\), one gets \(n\) values of \(q\) giving the \(n\) frequencies corresponding to the displacements at right angles to the plane of the undisturbed orbit.

Thomson proceeds then to calculate the frequencies for various numbers of particles.
Case of two electrons — \( n = 2 \). In this case one has the following six frequencies corresponding to the six degrees of freedom.

\[
0, \quad \omega, \quad \sqrt{\frac{\nu e^2}{m b^3}}, \quad \sqrt{\frac{\nu e^2}{m b^3}} - \omega, \quad \sqrt{\frac{\nu e^2}{m b^3}} + \omega, \quad \sqrt{\frac{3\nu e^2}{m b^3}} + \omega^2.
\]

If the particles are not rotating, \( \omega = 0 \) and the six frequencies degenerate (if we use the term in quantum mechanics) to the three frequencies, the first two, the following three and the last, given by

\[
0, \quad \sqrt{\frac{\nu e^2}{m b^3}}, \quad \sqrt{\frac{3\nu e^2}{m b^3}}.
\]

The rotation separates these frequencies except the last into two or three.

Case of three electrons — \( n = 3 \). Similar to the case of \( n = 2 \).

Case of four electrons — \( n = 4 \). In this case the oscillations in the plane of the undisturbed orbit remain the same as in the above cases, but about the oscillations at right angles to this plane, \( q^2 < 0 \) unless

\[
\omega^2 > \frac{4\sqrt{2} - 2\nu e^2}{8\sqrt{2} m b^3} > 325 \frac{\nu e^2}{m b^3},
\]

and the equilibrium is unstable, the four electrons then arranging themselves at the corner of a regular tetrahedron. If \( \omega \) is large enough to satisfy the above condition, the four electrons will be in steady motion in one plane at the corners of a square.

Case of five electrons — \( n = 5 \). This case is similar to the case of \( n = 4 \). For the oscillations at right angles to the plane of the undisturbed plane to be stable, it must be valid that

\[
\omega^2 > \frac{8.42\nu e^2}{19.42 m b^3} > 0.433 \frac{\nu e^2}{m b^3}.
\]

Case of six electrons — \( n = 6 \). To see the instability in this case, let us take the frequency equation for \( k = 3 \)

\[
\left(-\frac{14 - 8\sqrt{3}}{8} \cdot \frac{e^2}{a^2} - m q^2\right) \left(58 \frac{e^2}{8a^3} - m q^2\right) = 4m^2 \omega^2 q^2.
\]

As \( 14 - 8\sqrt{3} > 0 \), we see that one of the roots of this equation for \( q^2 \) is negative, so that the steady motion of six electrons in a ring is unstable, however rapid the rotation.
The motion can be made stable, however, by the introduction of an electron at the center. If a negative charge equal to that of \( p \) electrons is at the center of the ring, the radial force it exerts on the \( s \)th electron is \( pe^2/(a + \rho)^2 \), or \( pe^2/a^2 + 2pe^2\rho/a^3 \). When this term is put into the expression for the radial force, the frequency equation becomes

\[
\left( \frac{3}{4} \frac{e^2}{a^3} S_0 + \frac{3pe^2}{a^3} + L_k - L_0 - mq^2 \right) (N_0 - N_k - mq^2) = (M_k - 2\mu\omega q)^2. \tag{133}
\]

Applying this equation to the case of \( n = 6 \), in which an electron is at the center and five make an equilateral pentagon, we have with \( p = 1 \) the frequency equation

\[
\left( \frac{10 + 8\sqrt{3}}{8} \frac{e^2}{a^3} - m q^2 \right) \left( \frac{58e^2}{8a^3} - m q^2 \right) = 4m^2\omega^2q^2.
\]

The roots of this equation in \( q^2 \) are both positive, so that \( q \) is real and the equilibrium is stable.

Let us now see the condition for stability for displacements at right angles to the plane of the orbit. The frequency equation for \( k = 3 \) is

\[
\frac{\nu e^2}{b^3} - \frac{pe^2}{a^3} - \frac{34e^2}{8a^3} - mq^2 = 0.
\]

If \( p = 1 \), the condition for \( q^2 \) to be positive is

\[
\frac{\nu e^2}{b^3} - \frac{e^2}{a^3} - \frac{34e^2}{8a^3} > 0.
\]

We have, however,

\[
\frac{\nu e^2}{b^3} = m\omega^2 + \frac{e^2}{a^3} + \frac{e^2}{4a^3} \left( 5 + \frac{4}{\sqrt{3}} \right),
\]

so that

\[
m\omega^2 > 0.46\frac{\nu e^2}{b^3}.
\]

Let us then see the stability of the electron at the center of the ring. If it is displaced through a distance \( z \) at right angles to the ring, the equation of its motion is

\[
m\frac{d^2z}{dt^2} = -\frac{\nu e^2}{b^3}z + \frac{6e^2}{a^3}z.
\]
Thus if the motion is stable $\nu \frac{e^2}{b^3} > 6 \frac{e^2}{a^3}$, so that it results that

$$m\omega^2 > 0.53 \frac{\nu e^2}{b^3}.$$ 

This value of $\omega^2$ is greater than that for the condition of stability of the ring.

If the central electron, instead of being in the plane of the ring, was on one side of the center of the sphere of positive electrification while the ring was on the other side, the rotation required to make the equilibrium of detached electron stable would be less than when it was in the plane of the ring. Thomson mentions that for equilibrium the distance of the detached electron from the center of the sphere must be six times the distance of the plane of the ring from that point.

**Conditions for the stability of rings containing more than six electrons.** Thomson proceeds with the calculation, and points out that a single electron in the center is sufficient to make rings of seven and eight electrons stable, one of the values of $q^2$ being, however, exceedingly small though positive in the latter case.

When the number of electrons exceeds eight the number of central electrons required to ensure stability increases very rapidly with the number of the electrons in the ring.

In the frequency Eq. (133), that is

$$\left(\frac{3e^2S_0}{4a^3} + \frac{3pe^2}{a^3} - (L_0 - L_k) - mq^2\right)\left(N_0 - N_k - mq^2\right) = \left(M_k - 2m\omega q\right)^2,$$

$N_0 - N_k$ is always positive and $M$ is small compared with $L$ and $N$. Hence this equation will have real roots if

$$\frac{3e^2S_0}{4a^3} + \frac{3pe^2}{a^3} - (L_0 - L_k) > 0.$$ 

The greatest value of $L_0 - L_k$ is got by putting $k = n/2$ when $n$ is even, and $k = (n - 1)/2$ when $n$ is odd. The condition that the values of $q$ should be real, i.e. that the equilibrium of the ring should be stable, is therefore

$$\frac{3pe^2}{a^3} > \begin{cases} 
(L_0 - L_{n/2}) - \frac{3e^2S_0}{4a^3} & \text{for even } n, \\
(L_0 - L_{(n-1)/2}) - \frac{3e^2S_0}{4a^3} & \text{for odd } n.
\end{cases}$$
From these equations the least value of \( p \) can be calculated that will make a ring of \( n \) electrons stable:

\[
n \cdots 5, 6, 7, 8, 9, 10, 15, 20, 30, 40
\]

\[
p \cdots 0, 1, 1, 2, 3, 15, 39, 101, 232.
\]

For large values of \( n \) the values of \( p \) are proportional to \( n^3 \).

When \( p > 1 \) the internal electrons cannot all be at the center of the sphere, and they will separate until their repulsions are balanced by the attraction of the positive electricity in the sphere. Thomson says that when there are two electrons, as when \( n = 9 \), these two will separate and will form a pair with the line joining them parallel to the plane of the ring.

As it can be approximately assumed that the pair of equal electrons exerts at external points the same force as a double charge placed at a point midway between them, the preceding theory will apply, and the system consisting of the ring of 9 and the pair of electrons will be in stable equilibrium. When \( n = 10 \), the internal electrons must be three in number, and they will arrange themselves at the corners of an equilateral triangle, the plane of the triangle being parallel to but not coincident with that of the ring.

For a ring of 12 electrons, 7 inside electrons are required. But 7 electrons, as is mentioned above, cannot form a single ring, and will arrange themselves as a ring of 6 with one at the center.

In this way Thomson concludes that when there are a large number of electrons in rapid rotation they will arrange themselves as follows — The electrons form a series of rings, the electrons in one ring being approximately in a plane at right angles to the axis of rotation, and the number of particles in the rings diminishing as the radius of the ring diminishes.

Thomson says that the arrangement of electrons described above is like that of Mayer's floating magnets. In the case of floating magnets, they are not in rotation but are constrained to be on the surface of water. Thomson thinks that in the case of electrons the rotation makes their arrangement stable for their motion at right angles to the plane of the ring.

**Construction of the periodic system**

On the basis of the aforementioned theory, Thomson treats the problems (1) how would the corpuscles arrange themselves in the sphere of positive electrification, and (2) what properties would this structure confer upon the atom.
Let us see in the first place the problem (1). The solution of this is indicated by the results he has just obtained, when the electrons are arranged in concentric rings.

However, when the electrons are not constrained to one plane, they will arrange themselves in a series of concentric shells. Also in this case, as in the case of the ring, a number of electrons cannot be stably arranged over the surface of a shell. The stability is obtained by introducing within the shell an approximate number of other electrons, too. About the arrangement in shells Thomson says — "The analytical and geometrical difficulties of the problem of the distribution of the corpuscles when they are arranged in shells are much greater than when they are arranged in rings, and I have not as yet succeeded in getting a general solution. We can see, however, that the same kind of properties will be associated with the shells as with the rings."

Thomson asserts then that the property of the periodic system is associated with the solution in the case of the rings. He says — "As our solution of this case enables us to give definite results, I shall confine myself to this case, and endeavour to show that the properties conferred on the atom by this ring structure are analogous in many respects to those possessed by the atoms of the chemical elements, and that in particular the properties of the atom will depend upon its atomic weight in a way very analogous to that expressed by the periodic law."

Let the number of electrons be \( N \). Thomson considers as the simplest case the one in which the number of rings is minimum, so that in each ring there are as nearly as possible as many electrons as it is possible for the electrons inside to hold in equilibrium. Let the number of internal electrons required to make the equilibrium of a ring of \( n \) electrons stable be \( f(n) \). \( f(n) \) is denoted by \( p \) in the preceding discussions, and the values of \( p = f(n) \) are given in the table on p. 150 for various values of \( n \). The number of electrons in the outer ring \( n_1 \) will then be determined by the condition that \( N - n_1 \), the number of electrons inside, must be just sufficient to keep the ring of electrons in equilibrium, i.e. by the equation

\[
N - n_1 = f(n_1).
\]

If the value of \( n_1 \) got from this equation is not an integer, its integral part must be taken. The number of electrons in the second ring \( n_2 \) is determined, since there must be \( N - n_1 - n_2 \) electrons inside, by the equation
The Formation of Atomic Models

\[ N - n_1 - n_2 = f(n_2). \]

Similarly, the equations for \( n_3, n_4, \ldots \) are obtained.

Thomson solves these equations by a graphical method as follows. Draw the graph whose abscissa is \( f(n) \) and whose ordinate is \( n \). From the values of \( f(n) \) given on p. 150 for a series of values of \( n \) the curve in this figure is constructed. By means of the curve the numbers of electrons \( n_1, n_2, \ldots \) of the successive rings can be found. On the axis of abscissa take such a point \( P \) as gives \( N = OP \). Through \( P \) draw \( PQ \) inclined at an angle of 135° to the horizontal axis, cutting the curve in \( Q \), and draw the ordinate \( QM \). Then, \[ n_1 = \text{The integral part of } QM, \] is the number of electrons in the first ring reckoned from the outside. For evidently \[ OM = f(QM), \] and, because \( PM = QM \) and \( OM = OP - PM = OP - QM, \) \[ OP - QM = f(QM), \] which represents the equation \( N - n_1 = f(n_1) \) given above. To get \( n_2 \) similarly, mark off the abscissa \( OP_1 = N - n_1 \) (if \( QM \) is an integer \( P_1 \) will coincide with \( M \)), then from \( P_1 \) draw \( P_1Q_1 \) parallel to \( PQ \) cutting the curve in \( Q_1 \). The integral part of \( Q_1M_1 \) will be the value of \( n_2 \). In similar manners the values of \( n_3, n_4, \ldots \) will be found. In this way the following table is obtained.
The following table shows the series of arrangement of electrons for which the outer ring consists of 20 electrons (and before and behind by one).

<table>
<thead>
<tr>
<th>Number of electrons</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
<th>55</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Number in} )</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>13</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>( N )</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>( \text{Successive rings} )</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n_2 )</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n_3 )</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n_4 )</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n_5 )</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

59 is the smallest number of electrons which can have an outer ring of 20 electrons, while when the number of electrons is greater than 67 the outer ring will contain more than 20 electrons.

Thomson now touches upon the atomic weight. He supposes here also that the mass of every atom is entirely due to the sum of those of the atomic electrons, as we have mentioned at the beginning of the present subsection. In his words, "we suppose that the mass of an atom is the masses of the corpuscles it contains, so that the atomic weight of an element is measured by the number of corpuscles".

Thomson classifies in the first place the various arrangements of the electrons in families, the grouping of the electrons in the various members of the family having certain features in common. These features are formulated as follows:

The group of 60 electrons

= The same rings of electrons as the group of 40

+ a ring of 20 electrons,
The group of 40 electrons

= The same rings of electrons as the group of 24

+ an additional ring outside,

The group of 24 electrons

= The same rings of electrons as the group of 11

+ an additional ring outside.

The various groups of atoms are thus divided into series such that each member of the series is derived from the preceding member by adding to it another ring of electrons. Thomson expects that atoms formed by a series of electrons of this kind should have many points of resemblance.

The vibrations of the electrons in an atom

To see the resemblance of atoms in a series of this kind, Thomson takes in the first place the vibrations of the electrons in an atom. He divides these into two sets:

(1) Those arising from the rotation of the electrons around their orbits — if all the electrons in an atom have the same angular velocity, the frequency of the vibrations produced by the rotation of the ring of electrons is proportional to the number of electrons in the ring. Thus in the spectrum of each element in the series there would be a series of frequencies bearing the same ratio to each other, the ratios of the frequencies being the ratios of the numbers in the various rings.

(2) Those arising from the displacement of the ring from its circular figure — if the distance of an electron in the outer ring from an electron in the collection of rings inside, the effect of the outer ring will not alter the vibrations of the latter fundamentally.

For these vibrations, the sequence of frequencies would hence present much the same features for the various elements in the series, and there would be in the spectrum corresponding groups of associated lines.
About the perturbation caused on a ring by the presence of other rings, Thomson describes in his lecture in 1903 as follows.\textsuperscript{90)}

Suppose that the lines in the spectrum of three electrons are as represented in the Fig. a below. The numbers under the lines represent the number of periods which coalesce at that line; i.e., regarding the periods as given by an equation with nine roots, he supposes that there is only one root giving the period corresponding to the line A, while corresponding to B there are two equal roots, three equal roots corresponding to C, one root to D, and two to E. The absolute values of the periods are determined by the charge on the electron, the size of the sphere of the atom and the size of the ring, as in the calculation given above, but the ratio of the periods is not affected by these quantities. When the three electrons, instead of being alone in the sphere, form but one out of several rings in it, the absolute values of the periods would generally be entirely different, but ratio between them would be much more persistent. Thus, Thomson says that “using the phraseology of the Planetary Theory, we may regard the motion of the three corpuscles as ‘disturbed’ by the other groups”.

\text{Fig. a}

When there is only the group of three electrons, there are several displacements which give the same period of vibration; for example, corresponding to the line C there are three displacements, all giving the same period. When, however, there are other groups present, these different displacements will no

\textsuperscript{90)} J. J. Thomson, “Electricity and Matter” [the same as Ref. 85) of the present chapter], p. 119.
The Formation of Atomic Models

longer be symmetrical with respect to these groups, so that the three periods will no longer be quite equal. Thus, in the spectrum, C would become a triplet instead of a singlet, while B and E would become doublets.

The spectrum would now resemble the Fig. b. The more groups there are surrounding the group of three the more will be the motion of the latter be disturbed, and the spectrum would become like the Fig. c. Thus, if the elements which contain this particular group of electrons are classified according to the periodic law, there would be in the spectra of these elements homologous series of lines, the distances between the components of the doublets and triplets increasing with the atomic weight. In conclusion, Thomson says that “the investigations of Rydberg, Runge and Paschen and Keysel have shown the existence in the spectra of elements of the same group series of lines having properties in many respects analogous to those we have described”.

Thomson could not of course derive more profound laws of spectral lines than that mentioned above, for the same reason as Rayleigh’s theory of oscillation did not succeed in explaining the laws of spectral lines, as we have described in Sec. 1, §3 of the present chapter. The vibration of the electrons arranged along a ring is nothing but an example of the problem of oscillation treated by Rayleigh of a system of many identical bodies.

About valences

In the paragraphs concerning the construction of the periodic system, it has been mentioned that a series of atoms is formed. Now, Thomson regards that, when the atom of the pth number is formed from that of the (p – 1)th by addition of a single ring, they form a series which would all be in the same vertical column of the periodic table.

The gradual change in the properties of the elements along one of the horizontal rows in Mendeleef’s arrangement is also illustrated by the properties of these groups of electrons. Thomson considers, as an example, the series of arrangements of electrons given on p. 153 in which the outer ring contains 20 electrons. An outer row of 20 electrons first occurs with \( N = 59 \). In this case of \( N = 59 \), the number of electrons inside is only just sufficient to make the outer ring stable. This ring will therefore be on the verge of instability, and when the electrons in this ring are displaced the restoring force will be small. Thus when this ring is subjected to an external force, one or more electrons may
easily be detached from it. Such an atom will acquire a positive charge, and would behave like the atom of a strongly electropositive element. In the atom of $N = 60$ the outer ring is more stable, because there is an additional electron inside it. The greater $N$ becomes, the less electropositive the atom becomes.

When the stability of the outer ring gets very great, it may be possible for one or more electrons to be on the surface of the atom without breaking up the ring, bringing an electronegative atom. At $N = 68$ a great change in the properties of the atom would occur, since the number of electrons in the outer ring increases to 21, making the atom strongly electropositive.

The atom of $N = 59$ is very liable to lose an electron, and the arrangement of the electrons left would be the same as that in the case of $N = 58$, which has an outer ring of 19 electrons. This ring is exceedingly stable so that no further electrons would escape from it, while the positive charge on the system due to the escape of the 59th electron would attract the surrounding electrons to become neutral.

From these considerations Thomson derives valences. His argument for the elements at both ends of the series is in some sense farfetched. It goes as follows. The group of $N = 60$ would be the most electropositive of the series. But this could only lose one electron, for if it lost two we should have $N = 58$ as when the group of $N = 59$ had lost one electron, and in this case the system would be even more likely than the other to attract external electrons because it would have a positive charge of two units. Thus the system of $N = 60$ would get charged with one, and it would act like the atom of a monovalent electropositive element. The group of $N = 61$ would not part with its electrons so readily, but it could afford to lose two, acting like the atom of a divalent electropositive element, and so on. The group of $N = 67$, as soon as it requires a negative charge, would lose it again, and the system, like the group of $N = 59$, would be incapable of being permanently charged with electricity, acting like the atom of an element of no valence.

In this way Thomson deduces the following valences.

<table>
<thead>
<tr>
<th>No. of electrons</th>
<th>59</th>
<th>60</th>
<th>61</th>
<th>62</th>
<th>63</th>
<th>64</th>
<th>65</th>
<th>66</th>
<th>67</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valences</td>
<td>+0</td>
<td>+1</td>
<td>+2</td>
<td>+3</td>
<td>+4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>-0</td>
</tr>
<tr>
<td></td>
<td>-8</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>+5</td>
<td>+6</td>
<td>+7</td>
<td>+8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Electropositive</th>
<th>Electronegative</th>
</tr>
</thead>
</table>
He says that this sequence of properties is very like that observed in the case of the elements, and there is in fact the following series of elements.

\begin{equation*}
\text{He} \quad \text{Li} \quad \text{Be} \quad \text{B} \quad \text{C} \quad \text{N} \quad \text{O} \quad \text{F} \quad \text{Ne} \\
\text{Ne} \quad \text{Na} \quad \text{Mg} \quad \text{Al} \quad \text{Si} \quad \text{P} \quad \text{S} \quad \text{Cl} \quad \text{Ar}
\end{equation*}

The first and last of each of these lines are of no valence.

The construction of the periodic system and the theory of valences are, seen from the standpoint of the quantum theory, treated first by Bohr, and accomplished in quantum mechanics. Thomson's theory of the construction of periodic system should be said to be interesting as a prototype of the present-day quantum theory.

From our present-day standpoint, Thomson's relation of the sequence of valences, or atomic numbers, to that of atomic weights is very doubtful. Since he supposes that the mass of an atom is due to the sum of the masses of the electrons in it, one unit of atomic weight should mean about 1,000 electrons. However, since the sequence of atomic numbers carries with it the sequence of atomic weights which is more rapid in rise, the series shown above, for example, should not originate from the increase in $N$ by one in sequence, but should be due to the addition of $N$ of 1,000 units in each step, so that his way described above of constructing the periodic system is very questionable. This question comes to disappear in 1907 when Thomson abandons the idea that the mass of an atom is due to the masses of electrons.

\textit{About chemical combination}

Thomson considers the case of mixing electropositive (e.p.) atoms with electronegative (e.n.) atoms. This mixing may result in the transference of electrons from the e.p. atoms to the e.n. ones. The e.n. atoms will thus get a charge of negative electricity, and the e.p. atoms one of positive. The oppositely charged atoms will attract each other, and a chemical compound of the e.p. and e.n. atoms will be formed.

Just as an uncharged conducting sphere will by electrostatic induction attract a charged particle in its neighbourhood, so an electron outside an atom will be attracted, even though the atom has not become positively charged by losing an electron. When the outside electron is dragged into the atom there will be diminution in the potential energy.
If the potential energy diminishes when an electron falls into an atom A from an atom B, a mixture of both atoms results in the formation of a compound such as $A_B$. When an electron gets outside an atom in such a way, the presence of a medium of great specific inductive capacity such as water, or contact with a metal such as platinum black, would greatly increase the chance of chemical combination. Thomson explains so the catalysis.

Corpuscular temperature

In "Electricity and Matter" Thomson introduces the very noticeable concept of corpuscular temperature. This resembles Bohr’s introduction of the concept of nuclear temperature to treat the problem of atomic nuclei.\(^{91}\) This may be due to the resemblance in some respects of Thomson’s atoms to atomic nuclei known nowadays. Corpuscular temperature is the mean kinetic energy of the electrons inside the atom. The atom will not be stable unless its corpuscular temperature is below a certain value.

Thomson mentions the distinction between corpuscular temperature and the molecular temperature, which is the mean kinetic energy due to the motion of the center of gravity of the atom, pointing out a problem very important in principle.\(^{92}\) That is about the invalidity of the law of equipartition for the corpuscular and molecular temperatures. Thomson says — “These temperatures are probably not in any close relationship with each other. They would be proportional to each other if the law known as the law of equipartition of energy among the various degrees of freedom of the atom were to apply. This law is, however, inconsistent with the physical properties of gases, and in the proof given of it in the kinetic theory of gases, no estimate is given of the time required to establish the state contemplated by the law; it may be that this time is so long that gases are never able to get into this state.” In other words, he attempts to get over the invalidity of the law of equipartition of energy, by means of the great length of the time to reach this state.

Thomson takes the case of two atoms A and B interacting with each other. Their corpuscular temperatures are assumed to be high, though not so high as to make the atoms unstable when apart. As A and B approach each other,

---

\(^{91}\)This point was noticed also in K. Amano, "The History of Quantum Mechanics".

under their mutual attractions, the potential energy will diminish and their kinetic energy will increase. This increase in the kinetic energy of the electrons in A and B will increase the tendency of the electrons to leave their atoms, and if the increase in the kinetic energy is considerable A and B may each lose one or more electrons. A and B will become positively charged, and will tend to separate under the repulsion of these charges. These positive charges will ultimately be neutralized at a distance by the recombination with electrons. Then, unless the corpuscular temperature after union is less than a certain limiting value, the union cannot be permanent, the complex formed being unstable and incapable of a permanent existence. In other words, the corpuscular temperatures of A and B make a condition for the formation of complex aggregates. Thomson considers that, if the molecular temperature is very high, combination may be prevented by the high relative velocity of A and B carrying them apart in spite of their mutual attraction. This is opposite to the present-day idea of chemical combination. He emphasizes, in addition, that we cannot secure the union merely by lowering the molecular temperature, i.e., by cooling the gas, unless the corpuscular temperature is reduced below a certain value.

A specific example taken by Thomson in relation to his theory under consideration is interesting, as it shows the way of thinking at that time. He says — "The reason, on this view, why the atoms of hydrogen present on the earth do not combine to form some other elements, even at the exceedingly low temperature at which hydrogen becomes liquid, is that even at this temperature the kinetic energy of the corpuscles inside the atom, i.e., the corpuscular temperature, is too great. It may be useful to repeat here what we stated before, that there is no very intimate connection between the corpuscular and molecular temperatures, and that we may reduce the latter almost to the absolute zero without greatly affecting the former."

Thomson proceeds then to discuss the problem that different atoms combine to form a new atom. As for the conclusion he states — "Thus, if we regard the systems containing different numbers of units as corresponding to the different chemical elements, then as the universe gets older elements of higher and higher atomic weight may be expected to appear. Their appearance, however, will not involve the annihilation of the elements of lower atomic weight. The number of atoms of the latter will, of course, diminish, since the heavier elements are by hypothesis built up of material furnished by the lighter." He considers that there is a continual fall in the corpuscular temperature of the
atoms through radiation, and that the atomic weight of the lightest element surviving will continually increase, unless there is disintegration of the heavier atoms. He states, thus, that "on this view, since hydrogen is the lightest known element and the atom of hydrogen contains about a thousand corpuscles, all aggregations of less than a thousand units have entered into combination and are no longer free".

About the problem of the law of energy treated before, and the problem of the specific heats of gases in addition, Thomson argues as follows — "If the kinetic energy arising from the motion of the corpuscles relative to the centre of gravity of the atom could by collisions be transformed into kinetic energy due to the motion of the atom as a whole, i.e., into molecular temperature, it would follow from the kinetic theory of gases, since the number of corpuscles in the atom is exceedingly large, that the specific heat of a gas at constant pressure would be very nearly equal to the specific heat at constant volume; whereas, as a matter of fact, in no gas is there any approach to equality in these specific heats. We conclude, therefore, that it is not by collisions that the kinetic energy of the corpuscles is diminished."

The above facts, that the law of equipartition of energy is not valid between the corpuscular and molecular temperatures, and that the specific heat at constant pressure and that at constant volume of a gas take different values from each other, can be explained only on quantum theory. In other words, they cannot be explained on classical theory in which the electrons in an atom are considered to be in states of motion of continuous energy. The states of motion of the electrons must be of discontinuous energy, and the intervals of the discontinuous energies must in addition be so greater than the kinetic energies of the molecules that the states of motion of the electrons cannot change into those of higher energy by molecular collisions.

Explanation of radioactivity

According to Thomson's study of the stability of rings of electrons in his paper under consideration, the rings are stable when they are rotating with an angular velocity greater than a certain value, but become unstable when the velocity falls below this value. He takes also in this paper the case of electrons moving originally with velocities far exceeding the critical velocity. In consequence of the radiation from the moving electrons, their velocities will diminish. As we
have mentioned in §5 this radiation is very weak. These velocities will therefore reach the critical velocity after a long interval. Thomson says that, when the critical velocity is reached, there will be what is equivalent to an explosion of the electrons. They will move far away from their original positions, and their potential energy will decrease, while their kinetic energy will increase. "The kinetic energy gained in this way might be sufficient to carry the system out of the atom, and we should have, as in the case of radium, a part of the atom shot off". Thomson supposes that in consequence of the very slow dissipation of energy by radiation the life of the atom would be very long.

5. Model of the Atom with Nucleus

§1. Lenard's Dynamiden atom

It was made for the first time by P. Lenard\(^{93}\) to infer the structure of atoms from the scattering by atoms of charged particles of high velocity, or from the phenomenon of their freely passing through atoms. This study pioneered Rutherford's one. While Lenard made use of cathode rays, Rutherford applied \(\alpha\)-rays and determined finally the structure of atoms. In this sense Lenard's paper under consideration has a very important meaning. He measured the absorption of cathode rays of various velocities in material mediums. Lenard irradiated metallic plates with ultraviolet rays, and accelerated the electrons thus ejected from a plate by a voltage applied between the metallic plate and a metallic net put close to it. The accelerated electrons were let pass through a medium of various kinds, for example, a gas of various densities and pressures with a variable depth, and introduced into a metallic box to be measured of their intensity.

The law of mass absorption of cathode rays had been proposed by Lenard in his early study of cathode rays in 1895. This law asserted that for cathode rays of high velocity the absorptive power was proportional to the density of matter, irrespective of its state of aggregation and chemical composition.\(^{94}\) In other words, it stated that for cathode rays there was no distinction between two

\(^{93}\)P. Lenard, Ann. d. Phys. 12 (1903), 714.

\(^{94}\)Seen from the present, this law is of course wrong. As we shall see later, Lenard pointed out also this fact. The absorptive power is a complicated function of the atomic number \(Z\). In other words, it depends on the scattering in the field of atomic nuclei and that by atomic electrons.
mediums of the same mass. Lenard says in his paper cited just above that "the mass-law ... has, already before it was found that about the nature of cathode rays more than they would be processes in ether, suggested again and in the extreme the old hypothesis of the constitution of matter that various atoms of all the matter would be built of the same components in various numbers". Lenard refers to the fact that the theory of the constitution of matter, which Thomson mentioned at the end of his paper in 1897 on the study of cathode rays, used already such an idea, and notes that Thomson's idea has certain points common to, as well as different from, Lenard's one.

Lenard names these components Dynamiden. He says — "In order to be able to describe unambiguously these components, which in the following certain characteristics abstracted from the experience will be attributed to, I name them Dynamiden." Lenard supposes that every atom is made of such a number of Dynamiden that is proportional to its atomic weight. In other words, every material body is supposed to consist of a number of Dynamiden proportional to its weight. Two bodies of the same weight are thought therefore to differ exclusively in the different grouping of their Dynamiden of the same number. He supposes that it does not matter whether the body concerned is chemically simple or compound.

Lenard assumes that all the Dynamiden are exactly of the same weight as well as of the same mass, and in no occurring grouping they disturb each other in respect to their weights and masses. He ascribes the former mass-law to the equality of Dynamiden, and also to the lack of mutual disturbance with respect to their effect on cathode rays.

Lenard draws a knowledge of the space density of Dynamiden from the way of spreading of cathode rays in matter, and takes the value of $10^{-6} \sim 10^{-7}$ mm for the radius of atoms. He states that "it was closely put down through the manner of the spreading of cathode rays since the time when man studied to see quota moving on in matter", that atoms must show within their volumes still a structure made out of finer constituents, namely, Dynamiden. He points out that a quantity of rays moving rapidly enough can pass freely thousands of atoms, without changing the velocity in magnitude and direction. Even so, by going through matter a certain fraction of the quantity is fixed at atoms, in accordance with the absorptive power. Hence every Dynamide is to be given a certain absorbing cross section, in such a way that the quantity falling on the section is stopped, and the quantity passing near it is allowed to go through without any change in velocity.
The absorption of cathode rays by matter is represented by their intensity $I$ after they have passed through a depth $d$ of matter, which takes the form

$$I = Ae^{-ad},$$

where $a$ is the absorptive power. In the case of a gas, Lenard puts

$$a_0 = a/p,$$

$p$ being the pressure of the gas, and calls $a_0$ the specific absorptive power of the gas for a given velocity of cathode rays. He measures the values of $a_0$ for air and the gases of hydrogen, carbonic acid and argon. We cite here Lenard's table of the values of $a_0$.

<table>
<thead>
<tr>
<th>Velocity of cathode rays in Volt</th>
<th>Specific absorptive power $a_0$ at 1 mm Hg (in cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hydrogen</td>
</tr>
<tr>
<td>in ratio to light velocity</td>
<td></td>
</tr>
<tr>
<td>6 $\simeq 1/270$</td>
<td>44</td>
</tr>
<tr>
<td>30 $\simeq 1/120$</td>
<td>14.6</td>
</tr>
<tr>
<td>100 $\simeq 1/70$</td>
<td>6.01</td>
</tr>
<tr>
<td>1000 $\simeq 1/20$</td>
<td>1.2</td>
</tr>
<tr>
<td>4000 $\simeq 1/10$</td>
<td>0.19</td>
</tr>
<tr>
<td>$\simeq 30000$ $\simeq 1/3$</td>
<td>0.000062</td>
</tr>
<tr>
<td>— $\simeq 1/1$</td>
<td>0.0000006</td>
</tr>
</tbody>
</table>

About the values of $a_0$ Lenard considers, applying the kinetic theory of gases, that they represent the sum of the cross sections of the Dynamiden contained in the volume of 1 cm$^3$ at the pressure of 1 mm Hg.

From the measured values of $a_0$ shown in the table, it is seen that $a_0$ changes with the velocity of cathode rays, and therefore the cross section of a Dynamide depends on the velocity of cathode rays. Lenard concludes, in consideration of such an effect on cathode rays from the side of their character, that each Dynamide should be considered a field of electric force, and hence its cross section such a part of the field that is strong enough to stop cathode rays of a given velocity. About the nature of a Dynamide being a field of electric force, he says that "it could not seem doubtful any more that an electric character belongs essentially to every atom, since the spectra of elements were grasped according to Hertz as consequences of electric characteristic oscillations of atoms".
Lenard considers that there is a particular unpenetrable volume in a Dynamide, which he names the true cross section of Dynamide, and this must be supposed to be still smaller than the minimum cross section hitherto observed. He says that the minimum cross section measured is for cathode rays of a velocity almost close to the light velocity. In this case the absorptive power is very small. This fact means that the true cross section, or the true radius is quite small. For hydrogen gas one sees from the above table that the sum \( q \) of the cross sections of the Dynamiden contained in its volume of 1 cm\(^3\) at the pressure of 1 mm Hg is smaller than 0.0000006 cm\(^2\). Mayer calculated, on the other hand, the sum \( Q \) of the cross sections of hydrogen molecules contained in the same volume to be 13 cm\(^2\), from the internal friction of the gas.

Lenard puts

\[
Q = R^2 \pi N,
\]

\[
q = 2r^2 \pi Nz,
\]

where \( R \) is the molecular radius, \( r \) the true radius of a Dynamide, \( N \) the number of molecules in the volume concerned, and \( z \) the number of Dynamiden in a hydrogen atom. Taking

\[
q/Q = 2r^2 z/R^2 < 6 \times 10^{-7}/13,
\]

he gets

\[
r/R < 1.5 \times 10^{-4}/\sqrt{z}.
\]

Lenard assumes here the noticeable condition that

\[
z \geq 1.
\]

In other words, Lenard thinks that there is about one or more Dynamiden in a hydrogen atom. This means that he takes an assumption near to Prout's one, not supposing like Thomson that the mass of a hydrogen atom is due to the masses of about one thousand electrons. He has hence \( r < 1.5 \times 10^{-4} R \), from which he infers that

\[
r < 0.3 \times 10^{-10} \text{ mm}
\]

for the true radius of a Dynamide on the adoption of \( R \sim 0.2 \times 10^{-6} \text{ mm} \) due to Mayer. It is interesting that this is a value somewhat near the present-day
values of nuclear radii. Needless to say, the deduction of this value is based on a reasoning much different from those about the radius of atomic nuclei.

Now, for an atom of element of atomic weight $A$ the true volume of Dynamiden is $(4/3)r^3\pi Az$, whereas the volume of the atom is nearly equal to or greater than $(4/3)R^3\pi$ because all the elements hardly show atomic volumes smaller than that of hydrogen. The ratio of both the volumes is therefore smaller than $r^3Az/R^3$, or smaller than $3.4 \times 10^{-12}A/\sqrt{z}$ with the use of the value of $r/R$ given above. Since $A < 240$ and $z \geq 1$, this ratio is smaller than $10^{-9}$.

As all the material bodies have still rooms to move between them, it is valid all the more that in a volume $V$ of any body the volume available for Dynamiden is not more than $10^{-9}V$. Lenard says thus that the space in which there is a cubic meter of solid platinum, for example, is empty like the sky being passed through by light. We should note that the idea of the emptiness of the internal space of atoms like the present-day one was already at that time grasped with fair correctness. Lenard mentions — "The fact that the components considered of atoms are not effective at all by truly filling space, but almost exclusively by their electromagnetic field of force — together with their inner motion attributed —, justifies perhaps the name, which I have chosen for them on the assumption that the other meaning associated with this name formerly in the history of science has become foreign and hence would not be disturbing."

As Lenard supposes that the absorption of cathode rays is effected purely by the field of force of Dynamiden, he considers that this field of force can be known from the experiment on the absorption of cathode rays. This force of field must hence be learned from the above table. He infers in this way that this field of force is very strong from the absorption of cathode rays of 30,000 Volt.

Lenard considers that the discontinuation of the proportionality to mass with the transition to lowest velocities and the reduced growth of the absorption, can be understood in the way that the spheres of Dynamiden growing with the decrease in the velocity of rays come to overlap each other. He supposes that the degree of the mutual covering is determined by the ratio of the molecular volume to the molecular weight, whose values were given by Mayer to be 0.6, 0.1, 0.7 and 1.2 for hydrogen gas, air, argon gas and carbonic acid.

---

95) According to Lenard this name was given by Redenbacher. It was for the atom as a field of force in a unified theory of ether.
gas, respectively. Lenard thinks that the weaker absorption by argon gas of slow rays is for the same reason as the smallness of this ratio.

As the velocity of cathode rays tends to zero, the absorptive sphere of Dynamiden becomes equal to the total cross section of molecules. Lenard shows this by the following table.

<table>
<thead>
<tr>
<th>Gas</th>
<th>Molecular weight</th>
<th>Sum of molecular cross sections (cm⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(I) Collision with molecules of the same kind</td>
</tr>
<tr>
<td>H₂</td>
<td>2</td>
<td>12.3</td>
</tr>
<tr>
<td>Air</td>
<td>29</td>
<td>22.8</td>
</tr>
<tr>
<td>Ar</td>
<td>40</td>
<td>21.3</td>
</tr>
<tr>
<td>Co₂</td>
<td>44</td>
<td>33.5</td>
</tr>
</tbody>
</table>

As is seen from this table, for hydrogen gas the value in (II) is much larger than that in (I). This fact should be considered to show that a hydrogen atom exerts still a strong electric force quite outside its impenetrable space. Lenard infers that the average force in the space between the cross sections (I) and (II) should be of the order of 10⁹ Volt/cm for 6 Volt rays to be governed by both the cross sections.

Lenard points out that the mass-law is not valid even for cathode rays of high velocity. In other words, the ratio between the absorptive power and density, as well as its variation with the velocity of cathode rays, is nearly the same but not exactly. So he is led to assume that Dynamiden of various kinds of atoms possess their specific differences.

Then, Lenard enters further into the problem of atoms. The Dynamiden as the next components of electrically neutral atoms must be likewise electrically neutral. As the centers of field of force, however, they must have electric charges in them. He assumes hence, as the simplest idea of a Dynamide, an electric dipole of a certain moment. By means of the difference in the moment, Lenard explains the specific difference mentioned above of Dynamiden in each kind of atom. With such an idea of the structure of a Dynamide, he arrives at the conclusion that the upper limit given above of the true radius of Dynamide represents also the upper limit of the sum of the true and unpenetrable volumes of a negative and a positive electric elementary quanta. About this Lenard says
that "the smallness of this limiting value may be emphasized besides the mass of the positive elementary quantum which seems large for many indications".

Lenard's supposition that Dynamiden would be electric dipoles to compose atoms agrees with Thomson's one shown in 1903 in "Electricity and Matter". The difference between them consists in that Lenard supposes, adopting Prout's idea, a great mass for the positive elementary quantum, while Thomson attributed almost all the mass of an atom to electrons.

Lenard supposes also rapid motions, and hence an amount of kinetic energy of the Dynamiden in an atom, in order to explain radioactivity. He tries to explain the photoelectric effect, too, of which he made a detailed study himself. He gave an explanation of the photoelectric effect in his excellent study in 1902 of this effect.\(^{96}\) In other words, Lenard considered that the energy of the electron coming out of an atom in the photoelectric effect was not supplied by light, but was liberated from the energy stored up in the atom. The light was supposed only to play the role of a trigger. In this way he tried to understand the contradiction between the photoelectric effect and the classical theory, that electrons of a definite energy come out of the electrode regardless of its distance from the source of light. This explanation was not able, however, to elucidate his own discovery, that the energy of electrons changes with the frequency of light.

Lenard's model of the atom was the forerunner of the model of the atom with nucleus, but was not such one as to give enough ideas for the motion of the particles in atoms.

§2. Nagaoka's Saturn-like atom

It should be mentioned that we, Japanese physicists, receive a very reassuring tradition from the greatly important work carried out by Nagaoka, a pioneer of physics in Japan, in the history of the formation of atomic model. Nagaoka's work was read on Dec. 5, 1903 and was published in the next year.\(^{97}\) The Nagaoka model has aspects very much like those of the present-day model of atoms. It is clear that Nagaoka's paper played an important role in Rutherford's study, from the fact that Nagaoka's paper was cited by Rutherford, as well as the fact that other than Nagaoka there was no physicist who treated

\(^{96}\) P. Lenard, Ann. d. Phys. 8 (1902), 149.

\(^{97}\) H. Nagoka, Phil. Mag. 7 (1904), 445; Tokyo Sugaku-Butusrigakkwai Kiji-Gaiyo, II, No. 7 (1904), 92.
sufficiently such a kind of model. It should be noted at least that about the model of the atom with nucleus Rutherford did not cite any paper other than Nagaoka’s.

Nagaoka calculated the condition of stability and the mode of oscillation about the model of an atom, which had a heavy nucleus of positive electricity and rings of electrons rotating round it. His way of treating the calculation was somewhat similar to Thomson’s. Both of them took the example of Maxwell’s study of the stability of the rings of Saturn. Nagaoka attached importance to the spectral law, in contrast to Thomson who laid stress on the construction of the periodic system. The title of Nagaoka’s paper itself reads “Kinetics of a System of Particles illustrating the Line and the Band Spectrum and the Phenomena of Radioactivity”. However, in the meaning that both Nagaoka’s and Thomson’s models come to Rayleigh’s problem of oscillation we have described in Sec. 1, §3 of the present chapter, the former model is essentially the same as the latter model.

Now, Nagaoka says in his paper under consideration — “Since the discovery of the regularity of spectral lines, the kinetics of a material system giving rise to spectral vibrations has been a favourite subject of discussion among physicists. The method of inquiry has been generally to find a system which will give rise to vibrations conformable to the formulae given by Balmer, by Kayser and Runge, and by Rydberg.” He notices in particular the fact that line spectra show specifically the Zeeman effect in a magnetic field, but about band spectra this effect is not found. Nagaoka proposes, inasmuch as the empirical formulae are still a matter of dispute, to discuss a system whose small oscillations accord qualitatively with the regularity observed in the spectra, and by which the influence of magnetic field on band and line spectra is explicable, instead of seeking to find a system whose modes of vibration are brought into complete harmony with the mentioned regularity.

He mentions that the system here considered is quasi-stable, will serve to illustrate a dynamical analogy of radioactivity, and is perhaps the most easily conceivable, although the actual arrangement in a chemical atom may present complexities which are far beyond the reach of mathematical treatment.

The system to be treated consists of a large number of particles of equal mass arranged in a circle at equal angular intervals and repelling each other with forces inversely proportional to the square of distance. At the center of the

---

99) The Zeeman effect of band spectra was discovered by Wood in 1905.
circle, there is a particle of large mass attracting the other particles according to the same law of force. If these repelling particles are revolving with nearly the same velocity round the attracting center, the system will generally remain stable for small disturbances, provided the attracting force is sufficiently strong. Nagaoka states that "the system differs from Saturnian system considered by Maxwell in having repelling particles instead of attracting satellites". His words "The investigations on cathode rays and radioactivity have shown that such a system is conceivable as an ideal atom", indicates perhaps Lenard’s study we have described in the last subsection. About the origin of Nagaoka’s idea of Saturn-like atom he himself says — "In his lectures on electrons, Sir Oliver Lodge calls attention to a Saturnian system which probably will be of the same arrangement as that above spoken of." As regards the important problem inherent in the system considered, Nagaoka mentions that "the objection to such a system of electrons is that the system must ultimately come to rest, in consequence of the exhaustion of energy by radiation, if the loss be not properly compensated". It may be noted, however, there is a possibility of avoiding this difficulty by means of the very weak intensity of radiation shown by Thomson’s calculation we have described in §5 in the last section.

Nagaoka begins with the discussion of the stability of the system. Denoting the distance of the \( k \)th particle from the center of mass by \( R_k \), the total energy of the system by \( U \), and the potential energy by \( V \), one can prove the equation

\[
\frac{d^2}{dt^2} \sum mR_k^2 = 4U - 2V,
\]

for the law of force varying as the inverse square of distance, (according to Jacobi). Nagaoka points out that, in order that \( 2U - V \) should be neither always positive nor always negative, but assume oscillating values. Let the distance between the particles \( k \) and \( l \) be \( r_{kl} \), the repulsion between them \( e^2/r_{kl} \), the radius vector of the disturbed circular orbit \( r_k \), the attraction between the particle \( k \) and the central particle \( Ee/r_k \), and the angular velocity round the center \( \omega_k \), then approximately

\[
2U - V = \sum_k m\omega_k^2R_k^2 + \frac{1}{2} \sum_k \sum_l \frac{e^2}{r_{kl}} - E \sum_k \frac{e}{r_k}.
\]

Remembering that \( m\omega_kR_k^2 = \gamma_k \) represents the angular moment of the particle \( k \), which remains constant, one obtains
From this one sees that \( 2U - V \) will assume oscillating values, sometimes exceeding and sometimes falling short of zero, when \( r_k \) and \( r_{kl} \) are subject to small disturbances, provided the quantities \( e \), \( E \) and \( \omega_k \gamma_k \) and the mean values of \( r_k, r_{kl} \) are such as to bring the condition

\[
2U - V = 0
\]

repeatedly. One sees thus that for this to be satisfied one must have

\[
E \gg e.
\]

In relation to this, we must note that Nagaoka supposes, as we shall describe later, that \( E \) is about of the order of 10,000 \( e \). In that case the total mass of electrons which must exist to neutralize \( E \) becomes considerable, so that the mass of the positive particle at the center will be almost unnoticeable. In particular, there would arise difficulties for light elements.

Now, the small oscillations of a particle take place in three modes, namely oscillations in the radial and normal directions to the plane of its orbit, and that corresponding to the disturbance of \( r_{kl} \) which give rise to condensation and rarefaction of the particles arranged in a ring. About them Nagaoka says, citing Maxwell, that “the oscillations of mutually attracting particles have already engaged the attention of Maxwell in his discussion of Saturn’s rings; the equations here to be considered are nearly of the same form as those given by Maxwell, but they can be conveniently deduced by means of Lagrange’s equation”.

Let the radius of the undisturbed orbit be \( a \) and the position of any two particles \((1)\) and \((2)\) subtending an angle \(2\theta\) at the center be given by polar coordinates \( r, \phi, z \) such that

\[
\begin{align*}
    r_1 &= a(1 + \rho_1), \\
    r_2 &= a(1 + \rho_2), \\
    \phi_1 &= s + \omega t + \sigma_1, \\
    \phi_2 &= s + \omega t + \sigma_2 + 2\theta, \\
    z_1 &= a\zeta, \\
    z_2 &= a\zeta.
\end{align*}
\]

\( \rho, \sigma \) and \( \zeta \) being the radial, angular and transversal displacements, respectively.

The potential due to particle \((2)\) at the point \((r_1, \phi_1, z_1)\) is given by

\[
2U - V = \sum_k \frac{\omega_k \gamma_k}{r_k^2} + \frac{1}{2} \sum_k \sum_l \frac{e^2}{r_{kl}} - E \sum_k \frac{e}{r_k}.
\]
\[ V_{12} = -\frac{e}{r_{12}} = -\frac{e}{\sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos(\phi_2 - \phi_1) + (z_2 - z_1)^2}}. \]

Since
\[ r_{12} = 2a \sin \frac{\phi_2 - \phi_1}{2} \left[ 1 + \rho_1 + \rho_2 + \rho_1 \rho_2 + \frac{(\rho_2 - \rho_1)^2}{4 \sin^2 \theta} + \frac{(\zeta_2 - \zeta_1)^2}{4 \sin^2 \theta} \right]^{\frac{1}{2}}, \]
one finds by expansion
\[
\frac{1}{r_{12}} = \frac{1}{2a \sin \theta} \left\{ 1 - \frac{\rho_1 + \rho_2 + \rho_1 \rho_2}{2} + \frac{3}{8} (\rho_1 + \rho_2)^2 - \frac{(\rho_2 - \rho_1)^2}{8 \sin^2 \theta} - \frac{(\zeta_2 - \zeta_1)^2}{8 \sin^2 \theta} \right\}
\times \left\{ 1 - \frac{\sigma_2 - \sigma_1}{2} \cot \theta + \frac{(\sigma_2 - \sigma_1)^2}{8} + \frac{(\sigma_2 - \sigma_1)^2}{4 \cot^2 \theta} \right\}. \tag{136}
\]

Differentiating \( V_{12} \) and neglecting small quantities of second order one gets
\[
\begin{align*}
\frac{\partial V_{12}}{\partial r_1} &= -\frac{\partial V_{12}}{a \partial \rho_1} = \frac{e}{4a^2 \sin^2 \theta} \\
&\quad \times \left\{ 1 - \frac{\sigma_2 - \sigma_1}{2} \cot \theta - \frac{\rho_2 - \rho_1}{2} \cot^2 \theta - (\rho_2 + \rho_1) \right\},
\frac{\partial V_{12}}{r_1 \partial \phi} &= -\frac{\partial V_{12}}{a(1 + \rho_1) \partial \sigma_1} = \frac{e \cos \theta}{4a^2 \sin^2 \theta} \\
&\quad \times \left\{ 1 - \frac{3\rho_1 + \rho_2}{2} - \frac{\sigma_2 - \sigma_1}{2} (\tan \theta + 2 \cot \theta) \right\}, \tag{137}
\frac{\partial V_{12}}{\partial z_1} &= -\frac{\partial V_{12}}{a \partial \zeta_1} = \frac{e(\zeta_2 - \zeta_1)}{8a^2 \sin^2 \theta}.
\end{align*}
\]

On the supposition that the disturbances propagate round the ring in waves with the velocities \( n/h \) and \( n'/h \) for \( \rho, \sigma \) and \( \zeta \), respectively, one has
\[
\begin{align*}
\rho_1 &= A \cos(nt + \alpha + hs), \quad \rho_2 = A \cos(nt + \alpha + 2h\theta + hs), \\
\sigma_1 &= B \sin(nt + \alpha + hs), \quad \sigma_2 = B \sin(nt + \alpha + 2h\theta + hs), \tag{138}
\zeta_1 &= C \cos(n't + \gamma + hs), \quad \zeta_2 = C \cos(n't + \gamma + 2h\theta + hs),
\end{align*}
\]
where \( h = 1, 2, 3, \ldots \) \( h \) represents the mode of oscillation. Though Nagaoka does not mention much about the condition on this number, it is a number whose maximum value is determined by the number of particles, as in
Rayleigh’s and Thomson’s papers we have described in Secs. 1. §3 and 4. §6, respectively, of the present chapter. Since the case of a very large number of particles is treated in Nagaoka’s theory, a large enough number of spectral lines are supposed to arise.

Putting for simplicity

\[ u = nt + \alpha + hs, \quad u' = n't + \gamma + hs \]

one has

\[ \rho_2 - \rho_1 = A \sin u \sin 2h\theta + 2A \cos u \sin^2 h\theta, \]
\[ \sigma_2 - \sigma_1 = B \cos u \sin 2h\theta - 2B \sin u \sin^2 h\theta. \]

If the number of particles in a ring is denoted by \( \nu \), one sees that

\[ \theta = \frac{\pi}{\nu}, \frac{2\pi}{\nu}, \frac{3\pi}{\nu}, \ldots, \frac{(\nu - 1)\pi}{\nu}. \]

Let the following quantities be defined.\(^{100}\)

\[
L = \sum \left( \frac{1}{2} \frac{\sin h\theta \cos \theta}{\sin^2 \theta} - \frac{\cos^2 h\theta}{\sin \theta} \right),
\]
\[
M = \sum \frac{\sin 2h\theta \cos \theta}{4\sin^2 \theta},
\]
\[
N = \sum \frac{\sin^2 h\theta \cos^2 \theta}{\sin^3 \theta} + \frac{1}{2} \frac{\sin^2 h\theta}{\sin \theta},
\]
\[
J = \sum \frac{\sin^2 h\theta}{2\sin^3 \theta},
\]
\[
K = \sum \frac{1}{2\sin \theta}.
\]

\(^{100}\) G. A. Schott wrote a paper [Phil. Mag. 8 (1904), 384] criticizing Nagaoka’s theory, to which Nagaoka gave a refutation. In Schott’s paper Nagaoka’s mistake in the calculation in his paper was corrected as follows.

\[
L = \sum \left( \frac{1}{2} \frac{\sin^2 h\theta \cos^2 \theta}{\sin^3 \theta} - \frac{\cos^2 h\theta}{\sin \theta} \right).
\]
In these equations $\sum$ means the summation over the values of $\theta$. The potential at the point (1) due to the other particle is $\sum V_{12}$. If the suffix in $r_1, \phi_1, z_1$ is suppressed, the force at (1) is\textsuperscript{101}

$$
\begin{align*}
\frac{\partial V}{\partial r} &= \sum \frac{\partial V_{12}}{\partial r_1} = \frac{e}{2a^2} \{K + (LA - MB) \cos u\}, \\
\frac{\partial V}{r \partial \phi} &= \sum \frac{\partial V_{12}}{r_1 \partial \phi_1} = \frac{e}{2a^2} (MA + NB) \sin u, \\
\frac{\partial V}{\partial z} &= \sum \frac{\partial V_{12}}{\partial z_1} = -\frac{e}{2a^2} J \cos u'.
\end{align*}
$$

(140)

The kinetic energy of the particle at $(r, \phi, z)$ is given by

$$
T = \frac{1}{2} m \left\{ \left( \frac{dr}{dt} \right)^2 + r^2 \left( \frac{d\phi}{dt} \right)^2 + \left( \frac{dz}{dt} \right)^2 \right\}.
$$

The Lagrange equation is

$$
\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{r}} \right) - \frac{\partial T}{\partial r} = ma \frac{d^2 \rho}{dt^2} - ma \left( \omega^2 + \omega^2 \rho + 2\omega \frac{d\sigma}{dt} \right)
$$

$$
= -\frac{eE}{r^2} - \frac{e}{r} \frac{\partial V}{\partial r},
$$

with similar expressions for $\phi$ and $z$. Thus, one obtains the following equations of motion.

$$
\begin{align*}
\frac{d^2 \rho}{dt^2} - \omega^2 - \omega^2 \rho - 2\omega \frac{d\sigma}{dt} &= -\frac{e^2}{2ma^3} \{K + (LA - MB) \cos u\}, \\
\frac{d^2 \sigma}{dt^2} + 2\omega \frac{d\rho}{dt} &= -\frac{e^2}{2ma^2} (MA - NB) \sin u, \\
\frac{d^2 \zeta}{dt^2} &= -\frac{eE \zeta}{ma^2} + \frac{e^2}{2ma^2} J \cos u'.
\end{align*}
$$

(141)

\textsuperscript{101}Schott corrected the missing negative sign on the left-hand sides of Eq. (140), and pointed out that this caused Nagaoka's other mistakes. We write down Eq. (140) in the same form as Nagaoka's original equation.
From the last equation for $\zeta$ one finds as the frequency of transversal oscillation $n'$

$$n' = \pm \sqrt{S - \mu J}, \quad \text{(142)}$$

where

$$S = \frac{eE}{ma^3}, \quad \mu = \frac{e^2}{2ma^3}. \quad \text{(142)}$$

When $\nu$ is considerably large, $J$ in Eq. (139) can be expanded in a series of the form

$$J = A_j h^2 + B_j h^4 + \cdots, \quad \text{(143)}$$

so that one has

$$\pm n' = \omega_0' - a_j h^2 + b_j h^4 + \cdots, \quad \text{(143)}$$

where $\omega_0'$ is the principal constant term, and $a_j > 0$ and nearly proportional to $\nu^2/\omega_0'$. For small values of $h$, the values of $n'$ lie very near each other. Their intervals gradually become larger with increasing $h$, ultimately reaching their respective maximum values. The frequency decreases as $h$ increases. Nagaoka considers that these frequency lines give a band-spectrum. He says — “In fact, the above equation is but an extension of Deslandres’ empirical formula in a slightly altered form. It shows that the edges may lie either in the region of high or of low frequency. We shall afterwards find that the frequency is not affected by the external magnetic force, if the particles are supposed to be negative electrons.” Comparing the frequency lines with band-spectra, Nagaoka deduces the size of the ring. Namely, putting $e = 3.4 \times 10^{-10}$, $e/m = 2 \times 10^7$, $E = 15,000 \ e$, $\omega' = 2\pi u/\lambda = 2\pi(3 \times 10^{10}/0.0001) = 2 \times 10$, $u$ being the light velocity, he gets

$$2a = 6 \times 10^{-10} \ \text{cm},$$

which is about one tenth of the size of molecular sphere in the kinetic theory of gases. In one of his later papers,\textsuperscript{102} he takes $7 \times 10^{-8} \ \text{cm} \sim 10^7 \ \text{cm}$ for $a$.

In another paper,\textsuperscript{103} he takes a somewhat revised form for Eq. (143)

$$n' = a + bm^2 + cm^3 + dm^4 + em^5 + \cdots, \quad \text{(143)}$$

\textsuperscript{102}H. Nagaoka, Tokyo Sugaku-Buturigakkwai Kiji-Gaiyo II, No. 17 (1905), 240.
\textsuperscript{103}H. Nagaoka, ibid. II, No. 10 (1904), 129.
which includes terms up to \(m^5\), and considers that the result thereby deduced is in full agreement with Kayser’s and Runge’s formulae of band-spectra, making a comparison between it and the experimental knowledge of the band-spectrum of cyanide.

The other two equations in Eq. (141) are to determine the frequency \(n\) and angular velocity \(\omega\).

About these equations Schott pointed out Nagaoka’s mistake in calculation, and Nagaoka corrected them in his later paper.\(^{104}\) In this paper Nagaoka gives as the equations for the frequency of radial disturbance

\[
\omega^2 = S - \mu K ,
\]

\[
n^4 + \{\omega^2 + \mu (N - L - 2K)\}n^2 + 4\omega \mu Mn
\]

\[
- \mu N\{3\omega^2 + \mu (2K + L)\} - \mu^2 M^2 = 0 .
\]

It is seen that \(n^2 \sim \omega^2\) if \(\mu \ll \omega^2\). If the above equations are written in the form

\[
U(\omega + \varepsilon) = 0 ,
\]

\(\varepsilon\) is known to be

\[
\varepsilon = \frac{\mu}{2\omega} (4N - L - 2K \pm M) ,
\]

so that the first two roots of Eq. (144') are

\[
N_{1,2} = \pm \left\{\omega + \frac{\mu}{2\omega} (4N - L - 2K)\right\} + \frac{M}{2\omega} ,
\]

and the other two are

\[
n_{1,2} = \pm i\sqrt{3\mu N} + \frac{2\mu N}{\omega} .
\]

Nagaoka gets thus for the radial disturbance

\[
\rho = A_1 \cos(nt + hs + \alpha_1) + A_2 \cos(n_2t + hs + \alpha_2)
\]

\[
+ A_3 e^{\sqrt{3\mu N} t} \cos \left(\frac{2\mu M}{\omega} t + hs + \alpha_3\right)
\]

\(^{104}\)H. Nagaoka, ibid. II, No. 17 (1905), 240.
\[ + A_1 e^{-\sqrt{3\mu N}t} \cos \left( \frac{2\mu M}{\omega} t + hs + \alpha_4 \right). \tag{147} \]

A similar expression is obtained for the tangential disturbance.

If, in the above equation for \( n, N, L, K \) and \( M \) are expanded in the form

\[ a_0 + a_1 h^2 + a_2 h^4 + \cdots, \]

it results that \( n \) takes the form

\[ \pm n = \omega_0 + Ah^2 + Bh^4 \cdots, \tag{148} \]

which can also be written in the form

\[ \pm n = \frac{\omega}{\sqrt{a + bh^2 + ch^4 + \cdots}}. \]

Nagaoka notices now that waves of frequency \( n \) travel round the ring in opposite senses. The frequency increases as \( h \) is increased, and the spectral lines corresponding to these vibrations will gradually crowd together when \( h \) is large. He considers that the qualitative coincidence of this result with the line-spectrum is evident if \( h \) is not small. There is, however, an important discrepancy that when \( h \) is very great the intervals between successive lines begin to diverge, while in the empirical formulae the frequency tends ultimately to a limiting value as we have described in Sec. 1 of the present chapter on line-spectra. Nagaoka mentions that “it seems doubtful if very large values of \( h \) will ever be observed”.

He makes mention here of what is brought by the condition of the largeness of \( h \). In other words, Nagaoka says that the particles must be very small compared with the attracting center, in order for the ring not to collapse when disturbances corresponding to large values of \( h \) propagate round the ring.

He compares the line- and band-spectra so derived, to conclude that the successive difference in frequency is about nine times greater in the line-spectrum than in the band-spectrum.

Nagaoka discusses then the Zeeman effect. As we have mentioned above, at that time the Zeeman effect was known only for line-spectra, not being observed for band-spectra. Nagaoka considers, supposing that the moving particles are electrons, that the force acting on any one of them would be radial and equal to

\[ Ha \frac{d\phi}{dt} \approx eHa \omega, \]
when the magnetic field is perpendicular to the plane of the orbit. The equation of motion is therefore modified to give

$$\omega^2 = S - \mu K + \frac{eH}{m} \omega.$$  

It results therefore that

$$\begin{cases} 
\omega_1 = \sqrt{S - \mu K} + \frac{eH}{2m} , \\
\omega_2 = -\sqrt{S - \mu K} + \frac{eH}{2m} .
\end{cases}$$  \hspace{1cm} (149)

This means that there arise doublets, which are circularly polarized in opposite sense.

The magnetic field concerned does not affect, on the other hand, the transverse vibrations. But, when the field $H$ is parallel to the plane of the orbit, the force normal to it

$$eH \cos \phi \cdot \frac{d\phi}{dt}$$

will be operative. This force is oscillating, and its mean value is zero. Nagaoka concludes therefore that the Zeeman effect will not be produced in band-spectra.

Nagaoka notes Michelson and Lummer's finding that almost every spectral line is not a single line, and puts the cause of this to that the central particle may not actually be a point but may have a large size, and the other particles may be at different distances from the center.

Nagaoka considers that the doublets in the alkaline elements may be attributed to the Zeeman effect due to the magnetic field by other rings. But he mentions that it is extremely improbable that the field is so strong as to cause the observed separations. He considers also that, where there are many series of spectra, there must be the same number of rings of particles, and two neighbouring rings will be influenced so as to give rise to forced waves. He supposes cases in which the resonance due to the oscillation of other atoms may make the amplitude extremely large and ultimately tear the ring. The influence on the amplitude is also supposed to be one cause of the broadening of spectral lines.

About radioactivity — The solution (147) contains not only periodic terms but also a term with the factor $e^{\sqrt{3} \mu N t}$, so that if the disturbance persists the
ring will break. In this case, the particles will fly away with great velocities, and at the same time the central particle will participate in the same motion on account of the conservation of momentum. Nagaoka says that this explains $\alpha$- and $\beta$-radioactivity. Since $N \approx (h\nu/\pi)^2$ when $h$ is small, the greater $\nu$ is, namely the more massive the ring is, the greater the disturbance becomes, causing a quick disintegration of the atom. He considers that this is the reason why elements of great atomic weight accompany radioactivity.

It is not shown clearly, however, whether the central positive particle consists of $\alpha$-particles, or whether only a part of it flies away. Also, it is not clear how the other electrons behave after its flying away. If a ring disintegrates, much more electrons should fly away than those in $\beta$-radioactivity. In Nagaoka's theory these points remain obscure, as in Thomson's model of the atom.

Nagaoka discusses that the resonances due to forces, whose periods coincide with those of the rings, cause the destruction of the rings, and the modes of vibration which contribute to the instability of the system are those associated with the higher harmonics. He applies this way of reasoning to the explanation of the change of resistance observed in semi-conductors, and also to that of selenium and coherer. He mentions fluorescence, phosphorescence and luminescence as other examples of the forced oscillation.

§3. Arguments against Schott's critique

As we have mentioned above, Schott$^{105}$ corrected mistakes in Nagaoka's calculation, and criticized Nagaoka's theory of its not guaranteeing the stability of atoms.

In this paper, Schott takes the case of a ring consisting of $2p$ electrons. The most influential disturbing wave occurs when $h = p$. Inserting $\theta = \pi i/2p$ and $h\theta = \pi i/2$, where $i$ is an integer, into Eq. (139) in §2, and putting $a = 10^{-8}$ c.g.s. to have $\mu = e^2/2ma^2 = 10^{32}$, he finds

$$\frac{1}{2}K = 0.366 \times (2p) \log_{10}(2p), \quad \frac{1}{2}J = 0.017 \times (2p)^3.$$ 

For $2p = 10$, $K/2 = 3.6$ and $J/2 = 17$. Supposing the period of revolution to be the period of sodium light, he gets $\omega^2 = 10^{31}$. Unless the angular velocity is much greater, $\omega^2$ can be neglected in comparison with $\mu$ in the equation.

$^{105}$G. A. Schott, Phil. Mag. 8 (1904), 384.
\[ n^4 - \{\omega^2 + \mu(J - K)\}n^2 = \mu(2J - K)\{\omega^2 + 2n'^2 + \mu(3J - K)\} , \]

which comes from Eqs. (142), (144) and (144'). Similarly he neglects \( n'^2 \), getting

\[ n^2 = 67\mu \quad \text{or} \quad -37\mu . \]

The modules of the imaginary roots is about \( 6 \times 10^6 \), or about 120 times the number of revolution per second. These disturbances would therefore increase 2.7 times during \( 1/120 \) of a revolution. Schott concludes thus that it is impossible to have the stability of the system.

Schott discusses that, if the central charge is very large compared with the total charge of the ring, the ring can certainly be made stable, but in that case the remaining negative electrons necessary to make the whole system neutral are left out of account. He says that Nagaoka's system is not neutral electrically. He discusses also that it is not clear that the negative electrons outside the ring can be left out of account in treating the ring, for their charge is comparable in amount with the central charge. Assuming that they can, he shows \( \omega^2 \approx S \) from Eq. (144), and gets \( \omega^2a^2 = \nu c^2/ma \) with \( S = \nu c^2/ma^3 \). \( \omega a \) is the velocity of the ring. It must be less than the velocity of light. This gives \( \nu < 50,000 \). Next, Eq. (144') gives for \( -n^2 = 6\mu J \)

\[ \sqrt{-1} \frac{n}{\omega} = \sqrt{\frac{3J}{\nu}} . \]

For a ring of 10 electrons \( J \cong 34 \), and with \( \nu = 50,000 \) one has \( \sqrt{-1} n/\omega \cong 1/20 \). Schott says that the disturbance would increase 2.7-fold in about 3 revolutions, and the system is again far too unstable.

Against Schott's critique Nagaoka wrote the paper "Reply to Mr. Schott's Remark ...". In this paper, Nagaoka argues that the central charge is not supposed to be very large compared with the negative charges in the ring, the system investigated not being electrically neutral in Schott's sense. Nagaoka discusses that recent investigations show that every chemical atom is associated with numerous negative electrons, which in lightest elements amount to several hundred and in heavy elements reach hundred thousand. He considers that these electrons are not crowded together, and in making abstraction as to a hypothetical atom, one may conveniently assume that the central positive charge forms a neutral system with all the negative electrons, of which a small

\[ \text{[106] H. Nagaoka, Tokyo Sugaku-Butsurgakkwai Kiji-Gaiyo II, No. 11 (1904), 140.} \]
fraction will be found in a ring. He thinks also that the fact, that spectral lines
are not subject to a regular law and not expressible by a simple formula, will
probably find its explanation in perturbations due to stray electrons.

Nagaoka discusses this problem also in his next paper "The Structure of
an Atom", \(^{107}\) in which the correction of mistakes in his previous calculation is
given. He argues, about Schott's assumption of \(a = 10^{-8}\) cm, that \(a\) should be
taken larger because atoms are supposed to be flat. If \(V\) is the velocity of light,
the wavelength of light corresponding to the period \(T\) is given by \(\lambda = VT\). The
velocity of electrons describing a circular orbit of radius \(a\) is \(\omega a = V_0\), and if
they give rise to a light wave one has

\[ \lambda = \frac{a}{2\pi V_0}. \]

If it is assumed that \(\lambda = 10^{-4}\) cm and \(a = 10^{-8}\) cm, one has \(V/V_0 = 6 \times 10^4\),
which is apparently large. The limiting value \(V_0 = V\) is obtained by taking
\(a = 7 \times 10^{-8}\) cm with the number of negative electrons \(\nu = 200,000\). Nagaoka
considers that perhaps \(a = 10^{-7}\) cm is nearer the truth than \(a = 10^{-8}\) cm,
and consequently \(\omega^2 \gg \mu\).

Nagaoka states about the significance of his theoretical atom as follows.
"As the principal aim of the present investigation is to built up an ideal atom,
which will illustrate several physical properties of matter, it is inexpedient to
give from the outset several of the constants, found from measurements with
natural substances. I do not pretend that the system I propose represents a
chemical atom, but by investigating the properties of an ideal atom, we may
step by step aspire to a higher atom — the solution of the structure of natural
atoms." It should be said that these words are indeed suggestive.

In the paper under consideration Nagaoka discusses the vibration of the
heavy central body. He assumes \(E/M = 300\) (emu), \(M\) being its mass, and
finds for its vibration

\[ \Omega^2 = 3 \times 10^{23}. \]

This is about 1/300 of that of the D-line of Na, and corresponds to infrared
lines. He considers that the vibration will excite some vibration of the ring
of the electrons perpendicular to the plane of the orbit, giving rise to band
spectrum. He thinks that the vibration of the central positive particle in the
\(z\)-direction is caused by the combination with different atoms, and discusses

\(^{107}\)H. Nagaoka, \textit{ibid.} II, No. 17 (1905), 240.
that the band spectrum is a phenomenon inherent in molecules and not in atoms. In this connection, we can hardly be free from the impression, that the value of $E/M$ assumed by him is very incomprehensible in comparison with his previous supposition of $E = 200,000e$, because this value becomes consequently larger than the value of $e/m$ of an electron.

Nagaoka's theory of atoms as well as Thomson's has the difficulty common to classical theories that, while light is emitted by an atom through its vibration, the atom or the vibration loses energy by the amount of that of the emitted light, so that any monochromatic spectral line is impossible to occur, and wide continuous spectra are only expected. This difficulty comes to be solved by Bohr's theory of atoms.

6. Examinations of Both Models and the Determination of Atomic Model

§1. About the number of electrons in an atom-I

It is seen from the formation of the atomic models described above, that some confusion arose on account of the lack of any understanding of the substance of the positive electric charge, on the one hand, and the lack of any knowledge of the number of electrons in an atom, on the other hand. Only the atomic weight was the quantity that was known to be decisive of them. The atomic number itself, which plays an important role together with the atomic weight in the theory of atomic structure, had not yet gained any physical and structural significance, while the physical meaning of the order of atoms in the periodic table was considered to be possessed only by the atomic weight. The law of mass absorption of cathode rays (by Lenard) we have described before gave thus a proportional relation only to the atomic weight. The absorption of X-rays was considered also only in its correlation to the atomic weight.

About in 1905 various experiments were done to determine the number of atomic electrons. In particular Thomson paid attention to the problem of the number of atomic electrons, to elucidate the problem of atomic structure by synthesizing the experimental results. Thomson investigated this problem in 1906.\footnote{J. J. Thomson, Phil. Mag. 9 (1906), 769.} In the next year he also treated this problem in a book.\footnote{J. J. Thomson, "The Corpuscular Theory of Matter" (Archibald Consable and Co. Ltd., London, 1907), p. 142.}
this work, he says that “if we take the view that corpuscles are an essential constituent of the atom, one of the most fundamental questions to be answered is, how many corpuscles are there in an atom”.

Thomson makes use of three methods. They are (1) the intensity of secondary X-rays, (2) the absorption of cathode rays, and (3) the dispersion of light. According to him, all of them lead to the conclusion that the number of electrons in an atom is of the same order of magnitude as the atomic weight, and two of these methods show in addition that the ratio of the number of corpuscles in the atom to the atomic weight is the same for all elements. In relation to this, Thomson says that “the data at present available indicate that the number of corpuscles in the atom is equal to the atomic weight”, but “as the evidence is rather indirect and the data are not very numerous, further investigation is necessary before we can be sure of this equality; the evidence at present available seems, however, sufficient to establish the conclusion that the number of corpuscles is not greatly different from the atomic weight”. In other words, the methods used are very different and deal with widely separated physical phenomena. He thinks, however, that “although no one of the methods can be regarded as quite conclusive by itself, the evidence becomes very strong when we find that such different methods lead to practically identical results”.

(1) Method by means of secondary X-rays

This method is based on the determination of the proportion between the energy in a beam of primary X-rays passing through a gas and that in the secondary X-rays emitted by the gas. The theoretical formula of this process was given in Thomson’s book “Conduction of Electricity through Gases”. It is the Thomson scattering formula, which makes one of the fundamental equations in the classical electromagnetic theory. This formula gives the energy of the secondary radiation emitted per unit time per unit volume of a space containing corpuscles is equal to \((8\pi/3)Ne^4/m^2\) times the energy in the primary radiation passing through unit volume in that time, where \(N\) is the number of corpuscles per unit volume, \(e\) the charge on a corpuscle, and \(m\) its mass. This classical theory of scattering follows the following way of thinking. An electron emits a pulse while it changes its velocity. Now, if the electron is free, or if the force acting on it is weaker than that due to the pulse of the primary X-rays, the
change in the velocity of the electron takes place as long as the primary pulse passes over the electron. In this case, the thickness of the secondary pulse is equal to that of the primary pulse, and the penetrating power of the secondary rays is the same as that of the primary rays.

On the other hand, if an electron, after having been set in motion by the action of the primary rays, collides with another electron to receive a strong force, the quality of the secondary rays will be different. The secondary pulse will be thicker and easier to be absorbed, on account of the acceleration remaining in action after the passing through of the primary pulse. Further, if the primary pulse is thick, or if the electrons are close to each other, a thick secondary pulse will come out, because the pulses due to each electron are not separated but superposed. Thus, if the secondary rays are of the same quality as the primary pulse, it is possible to think that the primary pulse is thin enough, the electrons act individually, and the restoring force acting on the electron is weaker than the force due to the primary pulse, so that the Thomson formula holds good for free electrons.

The Thomson scattering is the one that is called the Compton scattering in quantum mechanics. The probability of the Compton scattering is exactly given by the Klein-Nishina formula. But it reduces to the Thomson formula when \( h\nu \ll mc^2 \), where \( h\nu \) is the energy of the primary X-ray and \( mc^2 \) the rest energy of the electron. To X-rays available at that time the Thomson formula applied almost correctly, provided that \( E_i \ll h\nu \), \( E_i \) being the ionizing energy to eject the electron.

Now, the study of secondary X-rays was done mainly by C. G. Barkla.\(^{110}\) He found that for light elements, hydrogen, air, H\(_2\)S, CO\(_2\) and CO\(_3\), the quality of the secondary X-rays is the same as that of the primary. He found also that the ratio of the energy of the secondary rays to that of the primary is independent of the quality of the primary rays, and is directly proportional to the density of the substance. Since this ratio is \((8\pi/3)Ne^4/m^2\),\(^{111}\) the density is known to be proportional to \(N\). And, because the density is given by the

\(^{110}\)C. G. Barkla, Phil. Mag. 5 (1903), 685; 7 (1904), 543.

\(^{111}\)We have already given, in Eq. (120) in the last subsection, the flux of the energy of the electromagnetic wave emitted from an electron in acceleration. Treating generally, let us calculate the energy flux, due to an electron moving at the point \(O\) with the acceleration \(\ddot{u}\), at the point \(P\) at the distance \(r\) from \(O\). If the component of \(\ddot{u}\) normal to \(OP\) is \(\ddot{u}_n\) and the angle between \(\ddot{u}\) and \(OP\) is \(\theta\), we have

\[
\ddot{u}_n = |\ddot{u}| \sin \theta .
\]
number of atoms in unit volume multiplied by the atomic weight, the number of electrons in an atom turns out to be proportional to the atomic weight. Thomson says that “the number of corpuscles in an atom of oxygen is sixteen times that in an atom of hydrogen, and so on”.

Barkla found that the energy of the secondary radiation from one cubic centimeter of air at atmospheric pressure is about 0.00025 times that of the primary radiation. Hence, one has

$$\frac{8\pi N e^4}{3 \ m^2} = 0.00025.$$  

The electric force is in the plane ($\hat{u}, \text{OP}$) and normal to OP, and the magnetic force is normal to the electric one. Let $\mathbf{E}$ and $\mathbf{H}$ be their respective fields. Then in the Gauss system of units

$$|\mathbf{E}| = |\mathbf{H}| = \frac{e\hat{u}_n}{c^2 r}.$$  

Because $\mathbf{E} \perp \mathbf{H}$, the density of energy flux $S$ is given by

$$|S| = \frac{c}{4\pi} |\mathbf{E}| \cdot |\mathbf{H}| = \frac{e^2 \hat{u}_n^2}{4\pi c^3 r^2} = \frac{e^2 |\hat{u}|^2}{4\pi c^3 r^2} \sin^2 \theta.$$  

Integrating this over the sphere of radius $r$, we get

$$S = \frac{2 e^2 |\hat{u}|^2}{3 c^3}.$$  

Let $\mathbf{E}_p$ be the electric field of the primary X-rays at the position of the electron. We have then

$$m\hat{u} = -e \mathbf{E}_p,$$

so that

$$S = \frac{2 e^4}{3 m^2 c^3} |\mathbf{E}_p|^2.$$  

The energy $P$ of the primary X-ray passing through unit area at the position O of the electron per unit time is, because $|\mathbf{H}_p| = |\mathbf{E}_p|$, 

$$P = \frac{c}{4\pi} |\mathbf{E}_p|^2.$$  

The ratio of the energy flux of the secondary rays to that of the primary is therefore

$$\frac{S}{P} = \frac{8\pi e^4}{3 m^2 c^4}.$$  

Being multiplied by the number $N$ of electrons, this gives the quantity desired. In the system of electromagnetic units, the last equation becomes

$$\frac{S}{P} = \frac{8\pi e^4}{3 m^2}.$$  

Because $e = 1.2 \times 10^{-20}$ and $e/m = 1.7 \times 10^7$, one finds

$$Ne = 10.$$ 

If $n$ is the number of molecules in a cubic centimeter of air at atmospheric pressure and $0^\circ$C, one knows

$$ne = 0.4.$$ 

Thus $N = 25n$, so that the molecules of air contain on the average about 25 corpuscles. Since the molecular weight of nitrogen is 28, the number of corpuscles in the atom of nitrogen turns out to be equal to the atomic weight of nitrogen.

The energy scattered by different gases is proportional to the density of the gas. Hence, if the number of corpuscles in an atom of any one substance is equal to the atomic weight of that substance, the number in an atom of any substance whatever must be equal to the atomic weight of that substance.

If some of the corpuscles are bound by forces so strong that they are practically rigidly attached to the atom, their acceleration would be very small, and give rise to very little second radiation. To this case Thomson applies his second method.

If the primary pulses are thicker than the atom, all the corpuscles in an atom ($p$ corpuscles, say) will behave as a single charged body. For a gas, in this case also the secondary pulses are of the same quality as that of the primary ones. Therefore, the corpuscles will act a body with a charge $pe$ and a mass $pm$, and the Thomson formula gives the expression $(8\pi/3)e^4p^2/m^2$. $p^2$ instead of $p$ will then be proportional to the atomic weight, the number of corpuscles not being proportional to it. Now, when a primary pulse of such great thickness is scattered by a solid or liquid, the secondary pulse would be very much thicker. Actually the secondary radiation from light liquids and solids, as well as that from gases, is of the same quality as that of the primary. Hence, the pulses in the primary radiation are concluded to be much thinner than the atom, so that $p$ but not $p^2$ is proportional to the atomic weight.

Barkla\textsuperscript{112)} made farther investigation on the quality of the secondary rays emitted by substances of larger atomic weight. He made clear that the secondary rays from a substance of large atomic weight are quite different from the primary X-rays, and are of an easily absorbable type. He showed

\textsuperscript{112}) C. G. Barkla, Phil. Mag. 9 (1906), 812.
that they are cathode rays, and that the secondary rays of the same penetrat-
ing type as the primary gradually diminish with the atomic weight. Barkla
made experiments about very many elements. He found that the secondary
rays have a little different quality from the primary rays, for carbon, paper,
wood, aluminum and sulfur, and have very weaker penetrating power than the
primary rays for calcium, iron, copper, zinc, tin, platinum and lead. In other
words, for these heavier elements the emission of cathode rays is predominant
and the secondary rays of large penetrating power are suppressed. Thomson’s
explanation of these facts is that in an atom of a heavy element the corpuscles
get so crowded that the X-ray pulses spread over more than one corpuscle at
once, and the direction of the electric force in the pulse is reversed between the
front and the rear of the pulse, to make the penetrating secondary radiation
diminish.

Barkla\textsuperscript{113)} made clear, in his later study in 1911, by measuring more pre-
cisely the scattering of X-rays, that the number of the scattering centers is just
equal to half the atomic weight. This gave for the first time the same correct
result as our present-day one. It is a very interesting fact that Rutherford
got contemporaneously with Barkla the same kind of result as to the positive
electric charge of the atomic nucleus.

Now, let the number of electrons in an atom be $Z$. Then $N = nZ$. The quantity

\[ S = \frac{8\pi}{3} \frac{e^4}{m^2} nZ \]

is called the absorptive coefficient due to scattering. $S/d$, $d$ being the density
of the substance concerned, is called the mass absorptive coefficient. Let the
atomic weight be $M$, with $M = 1$ for hydrogen, and the mass of the hydrogen
atom $m_H$, so that we have $d = nM = nm_H M/L$, where $L$ is the Loschmidt
number.

Barkla found the value

\[ \frac{S}{d} = 0.2, \]

for air, carbon and aluminum. For copper and silver he found larger values 0.4
and 1.5. Because for the lighter elements $(8\pi/3)e^4 L/m^2 = 0.40$, he got

\textsuperscript{113)} C. G. Barkla, Phil. Mag. 21 (1911), 684.
The Formation of Atomic Models

\[
\frac{Z}{M} = \frac{0.2}{0.40} = \frac{1}{2},
\]

which means that the number of electrons in an atom of these elements is half the atomic weight.

(2) Method by means of the absorption of cathode rays

If a cathode ray is traveling with a very high velocity through a collection of corpuscles, then when it passes close to one of the corpuscles it will be deflected; as the result of such deflections a bundle of cathode rays originally parallel to the \(x\)-axis will get more and more diffuse as they pass through the substance. The number passing in unit time through a unit area at right angles to the \(x\)-axis will therefore get smaller and smaller as the length of path of the rays increases. This was treated by Lenard, as we have already described, who supposed Dynamiden to constitute the atom and built a theory of their size and nature. Since the existence of electrons in the atom is now clear, Lenard's treatment must naturally be modified somewhat to take into account electrons as the scattering centers. In Thomson's model of the atom without nucleus the positive charge spreads uniformly over the whole atom contrary to Lenard's case, so that almost no scattering is caused by the positive charge. The electrons only act thus as the scattering bodies.

The amount of deflection experienced by the moving corpuscle depends on the force by which the corpuscles inside the atom are held in their positions of equilibrium. Since the solution of the problem when these forces are taken into account is difficult and complex, Thomson represents the action of these forces by increasing the mass of the corpuscles in the absorbing substance. The effect of a corpuscle that is held absolutely rigidly by some force acting on it, for example, is the same as if it were free from such force but had an infinite mass.

Let the inneratomic corpuscle be \(A\) with the mass \(M_1\), and the colliding corpuscle \(B\) with the mass \(M_2\). The velocity of \(B\) is assumed to be so great that \(A\) can be considered at rest in comparison with \(B\). Let the velocity of \(B\) before the collision be \(V\), and the perpendicular from \(A\) to \(V\) be \(b\). If \(AB = r\), the Coulomb force \(e^2/r^2\) acts between \(A\) and \(B\). The angle \(2\theta\), through which the direction of relative motion is deflected by the collision, is easily calculated. Then, if \(u, u'\) are the velocities of \(B\) parallel to the \(x\)-axis before and after the
collision, it is possible to obtain $u' - u$, which is now averaged with respect to the angle $\phi$ between the planes $(b, V)$ and $(V, x)$. If there are $N$ inneratomic corpuscles per unit volume, the number of collisions in which $b$ is between $b$ and $b + db$, made by a corpuscle B when it travels over a distance $\Delta x$, is $N\Delta x \cdot 2\pi bdb$. Let $U$ be the sum of the values of $u$ for the B corpuscles per unit volume, and $\Delta(U)$ the change in $U$ in the distance $\Delta x$. Then, $\Delta(U)$ is given by the equation

$$
\Delta(U) = -2UN \cdot \Delta x \frac{M_1}{M_1 + M_2} \int_0^{b'} \frac{2\pi bdb}{1 + \frac{b^2V^4}{e^4} \left( \frac{M_1M_2}{M_1 + M_2} \right)^2}.
$$

The upper limit $b'$ is determined by the condition that B comes into collision with the A corpuscles one at a time, so that the shortest distance between B and the corpuscle with which it comes into collision must be small compared with $a$, the distance between two inneratomic corpuscles. If $r$ is the shortest distance between A and B, we have

$$
1 - \frac{b^2}{r^2} = \frac{2e^2}{V^2r} \frac{M_1 + M_2}{M_1M_2}.
$$

Putting $r = a$, we see that $b'$ is of the order

$$
b' \sim a \left( \frac{1 - 2e^2}{V^2a} \frac{M_1 + M_2}{M_1M_2} \right)^{\frac{1}{2}}.
$$

The integration of Eq. (150) gives $dU/dx$ as a function of $b'$. The substitution of $b'$ given above in the integral results in

$$
\frac{d}{dx}(U) = -U \frac{4\pi Ne^4}{V^4} \frac{(M_1 + M_2)}{M_1M_2^2} \log \left( \frac{aV^2}{e^2} \frac{M_1M_2}{M_1 + M_2} - 1 \right).
$$

Thus $U$, the number of corpuscles crossing unit area in unit time, decreases as

$$
U = I_0 e^{-\lambda x},
$$

where $I_0$ is the value of $U$ at $x = 0$ and $\lambda$ is the absorptive coefficient given by

$$
\lambda = \frac{4\pi Ne^4}{V^4} \frac{(M_1 + M_2)}{M_1M_2^2} \log \left( \frac{aV^2}{e^2} \frac{M_1M_2}{M_1 + M_2} - 1 \right).
$$
If the corpuscles are quite free in the atom, \( M_1 = M_2 \) and we have

\[
\lambda = \frac{8\pi N e^4}{V^4 M_2^2} \log \left( \frac{1}{2} \frac{M_2 a V^2}{e^2} - 1 \right).
\]

If \( M_1 = \infty \), i.e. if the corpuscles are held fixed in the atom, we have

\[
\lambda = \frac{4\pi N e^4}{V^4 M_2^2} \log \left( \frac{M_2 a V^2}{e^2} - 1 \right).
\]

It should be noted that both the cases give the values of \( \lambda \) differing from each other only by the factor 2. Putting \( M_2 = m \) and expressing \( V \) in terms of the light velocity \( V_0 \), we have with the density of the substance \( \rho = (N/Z)Mm_H \) (\( Z \) being the number of corpuscles in the atom)

\[
\lambda \rho = 4\pi \frac{e^4}{m^2} \cdot \frac{eZ}{Mm_H} \times e \frac{V_0^4}{V^4} \log \left( \frac{aV^2 m}{V_0^2 e^2} - 1 \right).
\]

(152)

\( \lambda/\rho \) is almost constant independently of the nature of the substance. Hence \( Z \) is proportional to \( M \), because the logarithmic function varies slowly.

Bequerel gave \( V = 1.6 \times 10^{10} \) for \( \beta \)-rays from uranium, and Rutherford found \( \lambda/\rho = 7 \) for copper and silver with some uncertainty arising from \( a \) in the logarithmic factor. But Thomson considers that the uncertainty is not so great as to change the order of magnitude of this value of \( \lambda/\rho \). Substituting these values into Eq. (152), we have

\[
Z = M \frac{1.4}{\log \left( \frac{mV^2 a}{V_0^2 e^2} - 1 \right)}.
\]

Thus, \( Z \) is of the same order of magnitude as \( M \). In the book "Corpuscular Theory of Matter", Thomson applies the expression of \( \lambda \) for the case of \( M_1 = M_2 \), and makes use of Becker's measurements of cathode rays, that is, \( V = 10^{10} \) and \( \lambda/\rho = 1,200 \sim 2,000 \). Taking \( a \sim 10^{-8} \), Thomson obtains

\[
\frac{\lambda}{\rho} = 67 \frac{Z}{M} \log \left( \frac{10^5}{27} \right) \approx 556 \frac{Z}{M}.
\]

For the values of \( \lambda/\rho \) from 1,200 to 2,000, \( Z/M \) does not therefore take large values. This means that the number of corpuscles is of the same order of magnitude as the atomic weight.

Thomson says that this method supplements the preceding method, for on the former method corpuscles which are so firmly held that they are not moved...
Examinations of Both Models and the Determination of Atomic Model

by X-rays would not be accounted for. But he adds that the present method involves the assumption, that however small may be the distance between the corpuscles, the repulsion between them varies according to the Coulomb law.

(3) The method by means of the dispersion of light by gases

It was done for the first time by Drude\textsuperscript{114}) to determine the number of electrons in an atom, as famous Drude-Voigt's formula shows. Drude supposed electrons making linear oscillations of various frequencies to get the formula of the dispersion of light, by which he determined the number of electrons in an atom.

Thomson derives the formula of the dispersion of light in the following way. If an atom consisting of corpuscles dispersed through a sphere of uniform positive electrification is in the path of a wave of light, the electric field in the wave will displace the corpuscles. The motion of these corpuscles will produce a magnetic field, which will alter the velocity of light. Let $\xi_r, \eta_r, \zeta_r$ be the displacements parallel to the $x, y, z$-axis, respectively, of the $r$th corpuscle, and $x, y, z$ be the displacements of the center of the sphere of positive electrification. If $e$ is the charge of a corpuscle, and $E$ the charge of the sphere, the components of the electric force due to the displacement of the corpuscles $X', Y', Z'$ are given by

\begin{align}
X' &= \frac{4}{3} \pi N \left( Ex - \sum e \xi_r \right), \\
Y' &= \frac{4}{3} \pi N \left( Ey - \sum e \eta_r \right), \\
Z' &= \frac{4}{3} \pi N \left( Ez - \sum e \zeta_r \right).
\end{align}

The equations of motion for the corpuscles and the sphere are

\begin{align}
M \frac{d^2 x}{dt^2} &= (X + X')E - \frac{4}{3} \pi \rho e \sum (x - \xi_r), \\
m \frac{d^2 \sum \xi_r}{dt^2} &= -(X - X')E + \frac{4}{3} \pi \rho e \sum (x - \xi_r),
\end{align}

where \( m \) is the mass of a corpuscle, \( M \) that of the sphere and \( \rho \) the density of electricity of the sphere. If \( p \) is the frequency of the wave of light, all the quantities in question vary with the time as \( e^{ipt} \). Thus, we can solve Eq. (154) to obtain \( x \) and \( \sum \xi_r \). Substituting the solutions into Eq. (153), we get \( X' \). In the consequence of the motion of the charged corpuscles, the current is the polarization current \((K_0/4\pi)dX/dt\), \(K_0\) being the specific inductive capacity of the aether, plus the convection current \(N(Edx/dt - e\sum d\xi_r/dt)\). The total current determines the magnetic field, which gives by means of the Maxwell equations for \( X \)

\[
K_0 \frac{d^2 X}{dt^2} + \frac{3P}{1-P} \frac{d^2 X}{dt^2} = \frac{d^2 X}{dx^2} + \frac{d^2 X}{dy^2} + \frac{d^2 X}{dz^2},
\]

a wave equation, where

\[
P = \frac{(4/3)\pi N(mE^2 + MEe)}{(4/3)\pi \rho(Me + mE - mMp^2)}.
\]

Hence, if \( \mu \) is the refractive index, we have

\[
\mu^2 = 1 + \frac{3P}{1-P},
\]

or

\[
\frac{\mu^2 - 1}{\mu^2 + 2} = P.
\]

If \( p^2 \) is small, or the wavelength is long, we get

\[
\frac{\mu^2 - 1}{\mu^2 + 2} = \frac{NE}{\rho} \left(1 + \frac{M}{E} \frac{m}{e} \frac{p^2}{M + Zm} \frac{3E}{4\pi \rho}\right),
\]

because \( E = Ze \). Let \( a \) be the radius of the sphere of positive electricity. Then we have \( E = (4\pi/3)\rho a^3 \) and \( NE/\rho = (4\pi/3)Na^3 \). The latter is the volume of the atoms in 1 c.c., and is also the value of \((\mu^2 - 1)/(\mu^2 + 2)\) for \( p = 0 \), that is, for static electric fields. If this value is denoted by \( P_0 \) and the wavelength of light by \( \lambda \),

\[
\frac{\mu^2 - 1}{\mu^2 + 2} = P_0 + P_0 \frac{M}{E'} \frac{m}{e'} \frac{1}{(M + Zm)} \frac{3\pi}{\lambda^2},
\]
where $E'$ and $e'$ are the values of $E$ and $e$ in the system of electromagnetic units. Hence, it is seen that the dispersion tends to small as $Z$ and therefore $N$ become large.

Because of no available measurement done about the dispersion by monatomic molecules, Thomson makes use of Kelvin's finding that the dispersion by helium is of the same order of magnitude as that by diatomic molecules, to compare it with Ketteler's measurement done about the refraction index of hydrogen for light of various wavelengths. At the atmospheric pressure it is given by

$$\frac{\mu^2 - 1}{\mu^2 + 2} = \frac{1}{3} \left\{ 2.8014 \times 10^{-4} + \frac{2 \times 10^{-14}}{\lambda^2} \right\}.$$  

The comparison of this with the above equation gives

$$\frac{M}{E'} \frac{m}{e'} \frac{1}{(M + Zm)} = 6 \times 10^{-3}. \quad (155)$$

The values $m/e' = 1/1.7 \times 10^7$ and $Ne' = 0.8$ then give

$$\frac{M}{M + Zm} \frac{e'}{E'} = \frac{M}{M + Zm} \frac{1}{Z} = 1. \quad (156)$$

From this equation Thomson derives the following two conclusions:

1. $Z$ cannot be very much different from one,
2. $M$, the mass of the sphere of positive electricity, cannot be small compared with $Zm$, the total mass of the negatively charged particles.

His conclusion (1) means that the number of electrons in a hydrogen atom is not very much greater than one. It is important also to note his conclusion (2) saying that the number of electrons in a helium atom is not great following Rayleigh. This is to overthrow Thomson's earlier thought, and is in accord with the conclusion derived by Thomson from his study of positive rays round the time of his work under consideration.

In the above considerations it is assumed that the corpuscles are acted upon, when they are displaced, by the central force proportional to displacement due to the sphere of positive electricity, and the proportional coefficient is the same for all the corpuscles. Thomson says that this assumption does not hold good for a diatomic molecule, because the valence electron is acted upon by the forces of both the atoms with different coefficients of proportionality. He assigns thus,
The Formation of Atomic Models

according to Lorentz's theory, suitable inneratomic oscillations as Drude\textsuperscript{109}) did, without relating the atomic model in detail to inneratomic force. The refractive index $\mu$ for light of frequency $\nu$ due to a system of electrons with the charge $e$ and mass $m$, and having $p_1, p_2, p_3, \cdots$ for the frequencies of their oscillations about their positions of equilibrium, is then given by

$$\frac{\mu^2 - 1}{\mu^2 + 2} = \left( \frac{N_1 e^2}{p_1^2 - \nu^2} + \frac{N_2 e^2}{p_2^2 - \nu^2} + \cdots \right),$$

where $N_1$ is the number of electrons per unit volume having the frequency $p_1$, $N_2$ the number having the frequency $p_2$, and so on. $N$ is given by $N = N_1 + N_2 + \cdots$.

To the value of $(\mu^2 - 1)/(\mu^2 + 2)$ the oscillations with small $p$, make the largest contributions. In other words, the electrons for which the restoring force is weak are most easily influenced. Let $p_0$ be the frequency of such electrons, and $N_0$ the number of these electrons per unit volume. Then we have

$$\frac{\mu^2 - 1}{\mu^2 + 2} = \frac{N_0 e^2}{p_0^2 - \nu^2}.$$

Now, there is a number of substances for which the following relation is satisfied:

$$\frac{\mu^2 - 1}{\mu^2 + 2} = \frac{A}{p_0^2 - \nu^2}.$$

The values of $A$ are furnished by the experiments of Kettler and others. Comparing this with the preceding equation, we see that

$$A = \frac{N_0 e^2}{m},$$

from which $N_0$ is determined. This was done by Drude,\textsuperscript{114}) who found that the number so determined is greater than the number of atoms in that volume but not very much greater; it is very seldom, for example, as much as ten times greater. Drude also found that the greater the chemical valency of the atoms the greater is the number of the refracting systems per atom. This means that the valence electrons in an atom only take share in producing refraction, and the other electrons do not. Thomson concludes thus that “the optical
properties of other than monatomic gases are complicated by considerations which make them unsuitable for determining the total number of corpuscles in the atom”.

The number of spectral lines and that of inneratomic electrons

The reason that Thomson as well as Nagaoka supposed a very great number of inneratomic electrons is the existence of a great number of spectral lines. The spectral lines are supposed to be due to the oscillations of the electrons in the atom. Since for $Z$ electrons the degrees of freedom are $3Z$, the maximum number of different period of oscillation, i.e., of lines in the spectrum, is $3Z$. Thomson says thus that there may be such an objection to the preceding conclusion of the smallness of the number of inneratomic electrons, as “the number of corpuscles in the atom could not be less than one-third the number of lines in the spectrum, and this would for many elements be very much greater than the number representing the atomic weight”. If the Zeeman effect is taken into account, this number would be even greater. Thomson tries to explain this contradiction by considering that the majority of the spectral lines of an element do not arise from the atoms in the normal state, and that luminosity occurs when a gas is either traversed by an electric current or when it is raised to a high temperature, and in either case the gas is ionized, i.e., we have, in addition to the normal atoms, positively charged ions and negatively charged electrons. He gives also an explanation of the greatness of the number of periods by considering a system formed by a positive ion and a negative electron analogously to the solar system, or by supposing systems analogous to those in astromechanics such as a system of planets and others. Or, he supposes some central force in the atom that appears alternatively to be attractive and repulsive as the distance from the center increases, as in Boscovich’s idea of the atom.

§2. Origin of atomic mass

As we have mentioned before, Thomson thought round 1904 that the atomic mass was almost due to the atomic electrons, and hence that there were such a number of electrons in the atom as to correspond to its mass. One of the reasons of this supposition was his idea that all the masses were electromagnetic, so

---

that the positive charge with a great extension could have almost no mass, and only the electrons with a small radius could give rise to the atomic mass. However, the study of the number of electrons in an atom we have described in §1 made clear that this number is not several thousand or several ten thousand as was hitherto supposed, but is of the order of the number representing the atomic weight. Hence, it is found that the atomic mass is not due to the atomic electrons. Moreover, in the study of the number of atomic electrons by means of the dispersion of light, it is concluded that the mass possessed by the positive charge cannot be small compared with the total mass of the atomic electrons.

Together with these results, the mass of the positive charge is made clear in the study of positive rays. Positive rays were discovered in 1886 by Goldstein, and were named *Kanalstrahlen*. *Wien*\(^{116}\)) advanced the study of them to measure \(e/m\), and found that it took the largest value of \(10^4\) for hydrogen. Following this study, *Thomson*\(^{117}\)) made an excellent investigation.

From these things *Thomson*\(^{118}\)) concludes that the positive charge has a large mass. He says — "No positively electrified body has yet been found with a mass less than that of a hydrogen atom. We cannot, however, without further investigation infer from this that the mass of the unit charge of positive electricity is equal to the mass of the hydrogen atom, for all we know about the electrified system is, that the positive electricity is in excess by one unit over the negative electricity; any system containing \(n\) units of positive electricity and \((n - 1)\) corpuscles would satisfy this condition whatever might be the value of \(n\)." He thus mentions that before we can deduce any conclusions as to the mass of the unit of positive electricity we must know something about the number of corpuscles in the atom, to stress the significance of the study of the number of inneratomic electrons.

Referring to the result obtained in his book under consideration, namely, the one described above in §1 of the present section, that the number of corpuscles in an atom of any element is not much larger than the atomic weight, *Thomson* gives the following conclusions — "If this result is right, there cannot be a large number of corpuscles and therefore of units of positive electricity in an atom of hydrogen, and as the mass of a corpuscle is very small compared with that of an atom of hydrogen, it follows that only a small fraction of the

---


\(^{117}\)) *J. J. Thomson*, *Phil. Mag.* 8 (1907), 561.

\(^{118}\)) *J. J. Thomson*, "*The Corpuscular Theory of Matter*", p. 27.
mass of the atom can be due to the corpuscle. The bulk of the mass must be due to the positive electricity, and therefore the mass of unit positive charge must be large compared with that of the corpuscle, the unit negative charge.” Furthermore, Thomson found, as described in his book cited, two values of $10^4$ and $5 \times 10^3$ in the measurements of $e/m$ of positively electrified bodies, that is, two kinds of values of one-half unit and one unit, without finding any intermediate values. From this result Thomson concludes that “positive electricity is made up of units, which are independent of the nature of the substance which is the seat of the electrification”.

In spite of this conclusion, he does not throw away the idea that the positive electrification distributes uniformly inside the sphere of the atom. The distribution of positive electrification of such a radius can have only a trivial mass, if there is no mass other than electromagnetic ones. Thus Thomson mentions that this result shows that the mass is generated in two different ways; its part, the mass of an electron, is electric, and the other part of the atomic mass cannot but be thought to be mechanical. However, he does not distinguish these kinds of mass, and considers that it may be possible to take such a viewpoint as to dissolve the separation of these two natures of mass, suggesting a vortex theory like Kelvin’s early theory of vortex atoms. Thomson notices that a vortex moves not only with its tube but also with its surrounding liquid, and supposes thus that the largeness of the mass of a body of positive electricity compared with that of an electron shows the largeness of the quantity of ether carried by a positively electrified body in comparison with the quantity of ether carried by the tube of electric field and the seat of negative electricity.

§3. **Investigations of oscillations in atoms and of intensities of spectral lines**

As we have seen in Sec. 5, Nagaoka derived the modes of oscillation in atoms, and thereby tried to explain in particular line- and band-spectra. In his case, however, the application of the classical electromagnetic theory was not made to the oscillations in atoms, to derive the energies of radiations and compare them with the intensities of spectral lines. Such an investigation was done in 1907 by G. A. Schott.\(^{119}\)

At the beginning of this paper Schott says — “One of the most important problem of the Electron Theory of Matter is to account for the spectra by

\(^{119}\)G. A. Schott, Phil. Mag. 13 (1907), 190.
the several elements; the solution of this problem, rather than of any other, seems likely to lead to the construction of a working model of the atom." By Electron Theory of Matter he means any theory which assumes matter to consist of electrical charges, acting upon each other with electromagnetic forces only. He considers that it is necessary for stability to have both positive and negative charges, and one or both must be in more or less rapid motion. He considers also that since the atom is permanent, or very nearly so, the orbits of all the charges must be closed and of atomic dimensions, and that since it is electrically neutral, except when ionized, positive and negative charges must be present in equal amount. Schott states that the following investigation applies to both the Thomson and Nagaoka types of theory, except where special exception is made.

Schott points out that a single ring cannot be made to account for spectrum series; nevertheless its study throws much light on the three conditions to be satisfied:

(1) the electromagnetic waves emitted by the disturbed ring and received by a stationary observer must be of sufficient intensity to give observable lines;
(2) their frequencies must lie within the limits corresponding to the spectrum;
(3) they must be given a formula, such as that of Deslandres for bands, or those of Balmer, Rydberg, or Kayser and Runge for series.

Of these conditions, (3) is the most difficult to study, since it is required to solve more or less complicated frequency equations. Schott thus treats only (1) and (2). He says that, because they will be satisfied only for a number of the waves emitted by a single ring too small to account for the lines of even one series, it is useless to study (3). In his investigation, he uses a number of experimental data, namely, the intensities of strong and weak spectrum lines, the ionization in a gas giving a line spectrum, and the time for which a free ion exists on the average. He considers that, though the margin of error is large, a knowledge of the order of magnitude of these quantities is quite sufficient for the investigation.

In a previous paper, Schott examined the difficulty that a system of electrons in orbital motion may not have a definite structure at all, on account of the following two forces acting on an electron: (i) the reaction which the

\footnote{G. A. Schott, Phil. Mag. 12 (1906), 21.}
ether exerts on it in consequence of radiation, and (ii) a pull which the rest of
the ring exerts on it. He showed that changes produce no appreciable effect,
unless they involve a change in the order of the quantity which is changed.

Now, Schott shows first that stray electrons cannot exist in a permanent
atom. As we have mentioned in Sec. 4, §5, a calculation of this was made
by Thomson. Schott seeks the condition for this. Liénard, Abraham and
Schott gave independently the equation of the radiation from an electron in
acceleration. If \( u \) is the velocity of an electron and \( \beta = u/c \), \( c \) being the velocity
of light, the least value of the loss of energy per unit time of the electron by
radiation is given by putting \( \dot{\beta} = 0 \) to be
\[
R = \frac{2ce^2}{3(1 - \beta^2)^2} \cdot \frac{\beta^4}{\rho^2} = 2.4 \times 10^7 \frac{\beta^4}{(1 - \beta^2)^2} \text{ erg/sec},
\]
where \( \rho \) is the radius of the orbit taken to be of the order of \( 10^{-8} \) cm. The
comparison of this with the value of the kinetic energy of an \( \alpha \)- or \( \beta \)-particle
shows that the value of \( \beta \) is extremely small. Hence Schott considers that
its energy and the radius of the orbit may be calculated as if it obeyed the
ordinary laws of mechanics.

Its kinetic energy is \( E = c^2 m \beta^2 / 2 \). Let the positive charge be \( \nu e \). In the
model of Nagaoka type \( c^2 m \beta^2 / \rho \) is equal to \( \nu e^2 / \rho^2 \), while in the model of
Thomson type to \( \nu e^2 / b^3 \) if the radius of the sphere of uniform electrification
is \( b \). In any case \( c^2 m \beta^2 \geq \nu e^2 / b^3 \). Because \( R \approx 2ce^2 \beta^4 / 3p^2 \), it results that
\[
\frac{R}{E} = \frac{4e^2 \beta^2}{3cm^2} \geq \frac{4\nu e^2}{3c^3 m^2 b^2} \geq \nu \cdot 2 \times 10^9.
\]
In other words, a single electron cannot move in a circle of atomic radius for
any appreciable time. The electrons must therefore be arranged in groups. The
radiation from a group of electrons is of two types: (i) that due to the perma­
nent motion, and (ii) that due to disturbances produced by causes external to
the group.

The energy radiated by \( n \) elections moving uniformly in a ring at equal
intervals was given in Thomson's work we have described in Sec. 4, §5. Thom­
son's expression for it is valid, however, only when \( n \beta \) is small. Schott cal­
culates it without such a restriction, giving
\[
R = \frac{ce^2 \beta^2}{\rho^2} \sqrt{\frac{n^3 \gamma}{2\pi}} \exp \left[ n \left( \gamma - \frac{1}{2} \log \frac{1 + \gamma}{1 - \gamma} \right) \right]; \quad \gamma = \sqrt{1 - \beta^2}, \quad \beta < 1
\]
approximately for $n$ large. The reduction in the intensity of the radiation for $n$ large is due to the interference between the waves emitted by the several electrons in the ring. Each electron absorbs a great portion of the energy of the same type emitted by the remaining electrons in the ring. In other words, the reduction in the radiation from the electron is due to resonance. The reduction in radiation is by no means confined to the case of uniform circular motion. $n$ electrons moving in the same ellipse will interfere, if their eccentric angles increase at the same uniform rate, and are in arithmetic progression.

In the elementary explanation of the Zeeman effect, it is generally assumed that the D-lines of sodium can be attributed to the radiation of an electron moving in a circle. Schott proceeds to show that this is not so. For the D-lines $n\beta = 0.001,^{121}$ so that Thomson's expression suffices to give

\[
\begin{array}{lcccc}
 n & 1 & 2 & 3 \\
 R & 2.4 \times 10^{-5} & 3.6 \times 10^{-12} & 1.7 \times 10^{-19} \\
 R/E & 8.7 \times 10^{+7} & 26 & 1.7 \times 10^{-6} \\
\end{array}
\]

where $E = c^2 mn\beta^2 / 2$. These numbers show that the whole energy would be dissipated for 1 electron in one millionth sec., for 2 in 0.04 sec., for 3 in one fifth year. Schott compares this result with E. Wiedemann's measurement quoted in Drude's "Optik". It gives for the radiation from the D-lines $13.45 \times 10^{10}$ erg per sec per gram. The number of molecules at the standard state is $4 \times 10^9 / \text{c.c.}$ according Thomson's measurement, and the sodium molecule is supposed to be monatomic. Schott obtains hence $R = 3.9 \times 10^{-12}$ for each of the D-lines. This is of the order of the second case above, $n = 2$. In this case the atom can exist for at most 0.04 sec., whatever the conditions, for the radiation is continually taking place. Schott states that this is obviously impossible. He says in addition that apart from this it is known that spectrum lines vary very much in intensity and width with the conditions under which they are produced, which in itself is a sufficient proof that their energy is derived from external sources.

\footnote{Schott gives as the limit of spectral lines $\lambda = 10^{-5}$ cm and $\lambda = 8 \times 10^{-5}$ cm. With a times of revolution equal to the period of these extreme rays in a circle of atomic radius, an electron has velocities given by $\beta = 0.0063$ and $\beta = 0.0008$, respectively. In the case of $n$ electrons $n\beta$ takes these values.}
Schott considers then the disturbances possible in a ring of $n$ equidistant electrons in uniform circular motion. This case is, as we have mentioned at the beginning of the present subsection, what was treated first by Maxwell, and then by Thomson and by Nagaoka. These three treated only the modes of vibration, while Schott treats the energy of the radiation emitted through vibrations. Since every vibration accompanies its damping, he puts the displacements from steady motion of the $j$th electron to be given as

$$ (A, Be^{ia}, C) \exp \left[ -\kappa t + i \left( qt - k \frac{2\pi j}{n} \right) \right]. $$

\( \kappa \) represents the damping, \( q \) is the frequency seen from a system revolving with the mean angular velocity \( \omega \) of the ring, \( k \) represents the mode of vibration. In other words, the electrons at any instant lie on a curve of \( k \) waves, with \( 2k \) nodes and loops. These waves travel round the ring with angular velocity \( q/k \) forward, their angular velocity relative to a system at rest being \( q/k + \omega \). Hence the frequency seen from a system at rest is \( q + k\omega \). \( k \) can be so chosen as to lie between \( \pm n/2 \). Schott calls \( k \) the class. The forms of the vibration for the chief classes are shown in the following figure.

He is not concerned here with the forces which produce these vibrations, but only with the waves emitted. With the use of the Bessel function $J_{2m}(2l\beta)$, he

---

122) The displacements are the real parts of the components of this expression. The components refer to the tangential, radial and binormal displacements in the order given above.
obtains \( R \) in the form of an infinite series of harmonic terms. Here, \( m = k + sn \) and \( l = q/\omega + k + sn \), the frequency of the \( sth \) harmonic being \( q + (k + sn)\omega \). He points out that the fact, that a single vibration of the ring, corresponding to one degree of freedom of the ring, gives rise to an infinite series of harmonics, is due to the presence of ether which possesses an infinite number of degrees of freedom. If \( \lambda \) is the wavelength of the vibration considered, \( l\beta = 2\pi \rho /\lambda \). The distance between consecutive electrons in the ring is \( 2\pi \rho /n \). The amplitude \( A \) of the vibration can hardly be greater than a small fraction of this, that is, \( \sigma = nA/2\pi \rho \) is small. Schott gives the following values of \( R/\sigma^2 \) for various values of \( k \).

<table>
<thead>
<tr>
<th>Class, ( k )</th>
<th>0</th>
<th>( \pm 1 )</th>
<th>( \pm 2 )</th>
<th>( \pm 3 )</th>
<th>( \pm 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda = 6 \times 10^{-5} ) cm</td>
<td>0.0058</td>
<td>0.00029</td>
<td>( 4.8 \times 10^{-11} )</td>
<td>( 2.5 \times 10^{-18} )</td>
<td>( 6.4 \times 10^{-26} )</td>
</tr>
<tr>
<td>( \lambda = 3.5 \times 10^{-5} ) cm</td>
<td>0.0051</td>
<td>0.0025</td>
<td>( 1.2 \times 10^{-9} )</td>
<td>( 1.9 \times 10^{-16} )</td>
<td>( 1.2 \times 10^{-23} )</td>
</tr>
<tr>
<td>( \lambda = 2.10^{-5} ) cm</td>
<td>0.046</td>
<td>0.024</td>
<td>( 3.6 \times 10^{-8} )</td>
<td>( 1.7 \times 10^{-14} )</td>
<td>( 3.8 \times 10^{-21} )</td>
</tr>
</tbody>
</table>

It is frequently assumed that every disturbance of a system of electrons shows itself by a line in the spectrum of the system. Schott now shows that for a single permanent ring of electrons this is not the case, because many of the vibrations emitted are far too weak to affect the photographic plate appreciably.

He forms in the following way an estimate of the least amount of energy which will enable a vibration just to produce an impression on the plate. The photometric intensity is indicated by a scale number from 1 to 10. Each of the D lines is numbered 10. On the basis of Eder–Valenta’s and Kayser’s measurements Schott takes 1:400 for the ratio of the intensities of the scale numbers 1 and 10. This makes the energy of the weakest observable line \( 1 \times 10^{-14} \) erg per sec per atom, if the radiation in the D line is taken as \( 4 \times 10^{-12} \). He notes that this energy is underestimated, probably very much so for the ultraviolet.

Schott forms next some idea as to the values of \( \sigma \) for vibrations of different classes. For this purpose he considers the problem: — A ring of \( n + 1 \) equidistant electrons is in steady motion. An electron is suddenly expelled; required to find the subsequent motion. Now, in this case, (i) the nature of equidistance breaks, and (ii) the angular velocity and radius change from those for \( n + 1 \) electrons to those for \( n \) electrons. The influence due to (ii) is very
small compared with that due to (i). Let \( q, q' \) be the frequencies corresponding to classes \( k, -k \) respectively. As in the case of Saturnian rings considered by Maxwell, of the four frequencies for disturbances in the plane of the ring, two have tangential displacements large compared with the radial displacement, and the remaining two have those about twice as great.

For vibrations of class \( \pm k \), one has initially

<table>
<thead>
<tr>
<th>Class</th>
<th>( +k )</th>
<th>( -k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_0 )</td>
<td>( \frac{q'}{(n + 1)(q + q') \sin(k\pi/n)} )</td>
<td>( \frac{q}{(n + 1)(q + q') \sin(k\pi n)} )</td>
</tr>
<tr>
<td>Maximum kinetic energy</td>
<td>( \pi^2 m \rho^2 q^2 \sigma_0^2 / n )</td>
<td>( \pi^2 m \rho^2 q^2 \sigma_0^2 / n )</td>
</tr>
</tbody>
</table>

The ideal problem considered above differs from an actual case of ionization of a ring. The expulsion of the electron cannot really be instantaneous, but must be more or less gradual. The disturbance produced will therefore be less violent than in the ideal case, and the values of \( \sigma \) will be less. The vibrations of high classes will be weaker. The values of \( \sigma \) diminish with the time exponentially on account of the damping.

Schott examines now the effect of damping on radiation. He cites the conclusion of Stark's experiments on the Doppler effect of canal rays, that the line-spectra are emitted by positive ions produced from atoms by the expulsion of a negative electron, and the band-spectra in the recombination of the positive ion and negative electron. Schott compares thus a radiating gas to a system of a very large number of rings like those treated above. Suppose that on the average a ring remains ionized for \( T \) seconds. Further, suppose that the energy of an ionized ring due to a vibration of class \( k \) and amplitude \( \sigma \) is \( E = A\sigma^2 \), and the corresponding radiation \( R = B\sigma^2 \). One has

\[
\frac{dE}{dt} = -R,
\]

whence \( (dE/dt)/E = -R/E = -B/A \), a constant. Write \( B/A = \gamma \), then

\[
E = E_0 e^{-\gamma t}, \quad R = R_0 e^{-\gamma t},
\]

where \( t \) is reckoned from the instant of ionization.
The mean radiation from the gas per second per free ion is
\[ \frac{1}{T} \int_0^T Rdt = R_0 \frac{1 - e^{-\gamma T}}{\gamma T}. \]

When \( \gamma \) is very small as for the vibrations of high class, the last factor is practically unity and the radiation is \( R_0 \). When the damping is large as for the vibrations of low class, the last factor is nearly \( 1/\gamma T \), and the radiation is nearly \( R_0/\gamma T \), that is \( E_0/T \).

Thus the relative intensity of waves of high class and that of waves of low class are so different as to be used to distinguish both classes. The damping is given by \( \gamma = R_0/E_0 \), and for \( E_0 \) Schott takes the maximum kinetic energy, which is \( (\pi^2 m \rho^2 q^2/n)\sigma_0^2 \). Hence \( \gamma = (nR_0/\sigma_0^2)/\pi^2 m \rho^2 q^2 \), which may be written
\[ \gamma = \frac{nR_0/\sigma_0^2}{\pi^2 c^2 m (q\beta/\omega)^2}. \]

Then he tries to estimate the values of \( \gamma T \) for vibrations of different classes, in order to decide for which of these classes the intensity of the waves emitted is determined by the initial radiation, \( R_0 \), or by the initial energy, \( E_0 \). As we have mentioned above, a vibration of class \( k \) produces a series of harmonic waves, whose frequencies are \( q + k\omega + sn\omega \). The strongest of these, for \( k \) between \( \pm n/2 \), is given by \( s = 0 \). The others are extremely weak in proportion. If \( \lambda \) is the corresponding wavelength, one has \( q + k\omega = 2\pi c/\lambda \), and hence \( k\beta + q\beta/\omega = 2\pi \rho/\lambda \). For frequencies within the limits of the spectrum \( 2\pi \rho/\lambda \) is very small, so that \( q\beta/\omega \) is very nearly equal to \(-k\beta \). Therefore one finds for waves of corresponding frequency
\[ \gamma T = \frac{nT}{\pi^2 c^2 m \beta^2} \frac{R_0}{k^2 \sigma_0^2}. \]

The first factor depends only on the system of rings. The second factor depends on the class \( k \) of the vibration considered, and on \( R_0/\sigma_0^2 \) whose values are given in the above table. Schott supposed\(^{123} \) the expansion of electrons in order to avoid difficulties in the problem of spectral lines. Following this idea he advances now his discussions, to show that, except for small values of \( n \), \( n/\beta^2 \) may be taken to be 1,000. With this value \( n/\pi^2 c^2 m \beta^2 \) is of the order of \( 2 \times 10^8 \). The value of \( T \) was found by Thomson from the ionization of gases by X-rays

\(^{123}\)G. A. Schott, Phil. Mag. 12 (1906), 21.
to be of the order of $1/2$. The first factor in the last equation is therefore of the order of $10^8$.

With the use of the above table of $R/\sigma^2$, Schott shows that for $k = \pm 3$ or more $\gamma T \sim 10^{-7}$ or less, for $k = \pm 2$ $\gamma T \sim 1$ when $\lambda = 2 \times 10^{-5}$ cm and $\sim 0.001$ when $\lambda = 6 \times 10^{-5}$ cm, and for $k = 0, \pm 1 \gamma T$ is very large.

Thus the intensities of waves of classes $\pm 3$ and upwards are determined by radiation, those of waves of classes $0$ and $\pm 1$ by the initial energy, and those of classes $\pm 2$ mainly by $E_0/T$.

Schott compares these results with the intensities of spectral lines found in nature. The estimation made above of the unit amount of photometric intensity is based on the assumption, that every atom in the sodium flame at every instant shares in the radiation. If the radiation is assumed to be due only to the ions, the radiation per sec. per ion is greater in proportion.

He uses Starch's estimation that the ionization in the unstrained positive column of a vacuum tube is of the order of $10^{-5}$, and the ionization in a flame is much smaller. This value makes the radiation for a line of photometric intensity 1 at least $10^{-9}$ erg per sec per ion. This is the least value of $R_0$. Schott uses the table of values of $R/\sigma^2$ given above. The values of $\sigma_0$ are found from the table of equations given also above to be of the order of $1/k\pi$, for small values of $k$ and moderately large value of $n$. Applying these results to the case $k = +3$, he finds that $R \sim 2 \times 10^{-16}$ for $\lambda = 2 \times 10^{-5}$. This is only $1/5 \times 10^6$ of the value required for the faintest observable line. For greater wavelengths the value of $R$ is less still. Even if all the atoms are assumed to be ionized, the value of $R$ would be 50 times too small. It is thus concluded that a vibration of class $\pm 3$ or upwards is far weak to produce an observable spectral line.

For the vibrations of class $\pm 2$, the values of $R/\sigma^2$ are $6 \times 10^{-10}$, $3 \times 10^{-11}$ and $1.2 \times 10^{-12}$ erg per sec per ion for $\lambda = 2 \times 10^{-5}$, $3.5 \times 10^{-5}$ and $6 \times 10^{-5}$ cm respectively. Schott concludes that with an ionization $10^{-5}$ the system under consideration might give an extremely faint line in the ultraviolet, but could hardly give one in the visible spectrum. On the other hand, if all the atoms were ionized it could give a strong line in any part of the spectrum.

From the above discussions, it follows that observable spectral lines can only be produced by vibrations of classes $0$, $\pm 1$, and under certain circumstances by those of classes $\pm 2$. The relative intensities of these lines are given by $E_0/T$, and therefore depend mainly on $q$ and $q'$.
The vibrations of class 0 occur when the initial velocity and radius of the ring differ from those in the case of equilibrium, and the differences are the same for each electron of the ring. Let $\Delta \beta$ be the excess of the value of $\beta$ for $(n + 1)$ electrons above that for $n$ electrons in steady motion. Then the radiation is of the order $c^2 mn(\Delta \beta)^2 / 4T = (n\Delta \beta / 2\beta)^2 \times 10^{-9}$ erg per sec per ion. Schott calculates $\beta$ to show that the vibrations of class 0 for small values of $n$ give rise to an observable spectral line, even when the ionization is small, the line being stronger for $n$ smaller.

It is thus seen that the vibrations of classes 0, $\pm 1$, $\pm 2$ can produce lines sufficiently strong to be observable, but it is required that the frequency of any one of these lines is inside the limits of the spectrum. The condition for this is obtained from $k\beta + q\beta/\omega = 2\pi\rho/\lambda$ by putting $2\pi\rho/\lambda \geq 0.0008$ and $2\pi\rho/\lambda \leq 0.0063$, so that

for $q > 0$: $k = 0, -1$, and occasionally $-2$,

for $q < 0$: $k = 0, +1$, and occasionally $+2$.

That is, for a ring of electrons, each group corresponding to vibrations of the same type can give rise only to two, or at most three, observable lines.

With this result Schott criticizes Nagaoka's comparison of various modes of variation of a ring with line- and band-spectra. In other words, Schott argues that it is nonpermissible from the viewpoint of intensity to make the correspondence of each value of $k$ (h in Nagaoka's notation) with a line in a series. For example, of hydrogen the Balmer series has 29 lines, and a few series are known with scale numbers less 10. Even if all Nagaoka's 4 groups are supposed to combine to give a single series, only 12 lines at most will be obtained. For bands the difficulty is still greater. Furthermore, the theoretical intensities diminish very rapidly with the magnitude of $k$ larger than 2. Nothing like this is found in reality. Schott states that "this difficulty is conclusive against Nagaoka's view".

As a modification of Nagaoka's model, Schott discusses the case of higher powers of $\beta$. He points out that, in this case, each frequency equation is transcendental and therefore has an infinite number of roots, but for the ring only 18 lines are also found to be of sufficient intensity, so that the purpose is not attained.

From Schott's study described above, it results that the vibration of a ring of electrons cannot lead to a theory of spectral lines. Thus, the classical
electromagnetic theory is found to break down in the field of spectral lines. This will come to be solved for the first time by means of quantum theory.

§4. The determination by Rutherford

After the breakdown of the model of the atom with nucleus as well as of the model of the atom without nucleus in kinematic and dynamical problems, the final determination of the model of the atom was done by Rutherford.\footnote{124}{E. Rutherford, Phil. Mag. 21 (1911), 669.} Just before this determination, Thomson\footnote{125}{J. J. Thomson, Camb. Lit. & Phil. Soc. 15 (1910), 5.} was in reliance on the model of the atom without nucleus.

As we have described in the last section, Lenard tried to investigate the structure of atoms by making charged particles (cathode rays) pass through matter. Also some people tried later to clarify the structure of atoms with this method. They used as charged particles \(\alpha\)-rays as well as \(\beta\)-rays, of which \(\beta\)-rays suffer greater scattering on account of the smaller momentum and energy.

By the group of J. J. Thomson as its leader, a scattering with great deflection is supposed to result from a large number of scatterings with small deflection. Thomson's theory of scattering says as follows. When a negatively charged particle passes through an atom, it is deflected by (1) the repulsion of the electrons distributed through the atom, and (2) the attraction of the positive electricity in the atom. The deflection of the particle is supposed to be small. If \(\theta\) is the average deflection due to a single atom, the average deflection after a large number \(m\) of encounters is \(\sqrt{m}\theta\). As we have described in §1 of the present section, Thomson investigated the number of electrons in an atom by means of the scattering of cathode rays.

The accuracy of this theory of compound scattering was examined experimentally for \(\beta\)-rays by Crowther.\footnote{126}{J. A. Crowther, Proc. Roy. Soc. 84 (1910), 226.} He apparently confirmed Thomson's theory, concluding that the positive electricity distributed continuously throughout an atom, and the number of electrons in an atom was about three times its atomic weight.

However, Geiger and Marsden\footnote{127}{H. Geiger and E. Marsden, Proc. Roy. Soc. 82 (1909), 495.} made experiments on the scattering of \(\alpha\)-rays, and showed that some of the \(\alpha\)-particles suffered a deflection of more than a right angle at a single encounter. They found, for example, that about
1/20,000 of the incident α-particles were turned through an average angle of 90° in passing through a layer of gold-foil about 0.00004 cm thick, which was equivalent in stopping power of α-particles to 1.6 mm of air. Geiger\(^{128}\) showed later that the most probable angle of deflection for a beam of α-particles traversing a gold-foil of this thickness was about 0.87°.

Now, Rutherford points out in his paper under consideration, that these results are in complete opposition to Thomson's theory. In other words, Rutherford says that according to a calculation based on the theory of probability the chance of an α-particle being deflected through 90° is vanishingly small. He shows in addition that the distribution of α-particles for various angles of large deflection does not follow the probability law to be expected if such large deflections are made up of a large number of small deflections.

Thomson's assumption supposes a small scattering due to a single atom in opposition to the experiments by Geiger \textit{et al.}, and the structure assumed by Thomson for the atom does not admit of a very large deflection of an α-particle in traversing a single atom. Rutherford concludes therefore that for a large deflection to occur by a single atomic encounter a strong electric field is needed, or, in other words, the diameter of the sphere of positive electricity should be minute compared with the diameter of the sphere of influence of the atom.

In the experiments by Geiger it was necessary to observe a small numbers of α-particles. Rutherford stresses that this became possible by the finding by Geiger of scintillation method of counting individual α-particles.

After the consideration described above, Rutherford examines theoretically an atom of simple structure which is able to produce large deflections of an α-particle, and compares the deductions with the experimental data. It is an atom with nucleus, and has a character most common to that in the Nagaoka model in the domain of consideration necessary for Rutherford's theory.

Rutherford considers an atom which contains a charge \(±Ne\) at its center surrounded by a sphere of electrification containing a charge \(±N\varepsilon\) supposed uniformly distributed throughout a sphere of radius \(R\). \(e\) is the fundamental unit of charge equal to \(4.65 \times 10^{-10}\) e.s.u. He supposes that for distances larger than \(10^{-12}\) cm the central charge and also α-particle may be taken to be point charges. Rutherford mentions that, because the main deductions from his theory are independent of whether the central charge is positive or negative,

for convenience the sign will be assumed to be positive. He mentions also that
the question of the stability of the atom need not be considered, for this will
obviously depend upon the minute structure of the atom, and on the motion
of the constituent charged parts.

In the atom so assumed, the electric force $X$ and the potential $V$ at a
distance $r$ from its center are given by

\[ X = Ne \left( \frac{1}{r^2} - \frac{r}{R^2} \right), \]
\[ V = Ne \left( \frac{1}{r} - \frac{3}{2R} + \frac{r^2}{R^3} \right). \]

Suppose an $\alpha$-particle of mass $m$, velocity $u$ and charge $E$ shot directly towards
the center of the atom. It will be brought to rest at a distance $b$ from the center
given by

\[ \frac{1}{2} mu^2 = Ne \left( \frac{1}{b} - \frac{3}{2R} + \frac{b^3}{R^3} \right). \]

This $b$ is an important quantity in the theory. If $N = 100$ and $u = 2.09 \times 10^9$ cm/sec, $b = 3.4 \times 10^{-12}$ cm, while $R \approx 10^{-8}$ cm. As is seen from this value
of $b$, the $\alpha$-particle before being turned back penetrates so close to the central
charge, that the field due to the uniform distribution of negative electricity may
be neglected. In fact, for all deflections greater than a degree one may without
sensible error suppose the deflection to be due to the central charge alone.

If the velocity of the $\alpha$-particle is assumed not to be changed appreciable
by its passage through the atom, the path of the particle under the influence
of a repulsive force varying inversely as the square of the distance will be
a hyperbola with the center of the atom $S$ as the external focus (cf. Fig.).
Suppose the particle to enter the atom in the direction $PO$, and that the
direction of motion on escaping the atom is $OP'$. $A$ is the apse of the hyperbola,
and $p = SN$ is the perpendicular distance from the center on the direction of
initial motion. Let angle $POA = \theta$. 
With the initial velocity \( u \) of the particle and its velocity \( v \) at \( A \), one has from the conservation of angular momentum

\[
pu = SA \cdot v.
\]

From the conservation of energy

\[
\frac{1}{2} mu^2 = \frac{1}{2} mv^2 + \frac{NeE}{SA},
\]

\[
v^2 = u^2 \left( 1 - \frac{b}{SA} \right).
\]

Since the eccentricity is \( \sec \theta \),

\[
SA = SO + OA = p \cosec \theta (1 + \cos \theta)
\]

\[
= p \cot \theta / 2,
\]

\[
p^2 = SA(SA - b) = p \cot \theta / 2(p \cot \theta / 2 - b),
\]

\[
\therefore b = 2p \cot \theta.
\]

For the angle of deviation \( \phi = \pi - 2\theta \)

\[
\cot \frac{\phi}{2} = \frac{2p}{b},
\] (157)
which is valid also in the case of attractive force. This relation is illustrated in the following table.

<table>
<thead>
<tr>
<th>$p/b$</th>
<th>10</th>
<th>5</th>
<th>2</th>
<th>1</th>
<th>0.5</th>
<th>0.25</th>
<th>0.125</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>5°.7</td>
<td>11°.4</td>
<td>28°</td>
<td>53°</td>
<td>90°</td>
<td>127°</td>
<td>152°</td>
</tr>
</tbody>
</table>

Rutherford calculates next the probability of getting $\phi$ in a single collision. Suppose a beam of charged particles to fall on a thin screen of matter of thickness $t$. Let $n$ be the number of atoms in unit volume of the material. Then the number of collisions of the particle with the atom of radius $R$ is $\pi R^2 nt$ in the thickness $t$. The probability $m$ of entering an atom within a distance $p$ of its center is given by

$$m = \pi p^2 nt.$$  

The probability $dm$ of striking within radii $p$ and $p + dp$ is therefore

$$dm = 2\pi pnt dp = \frac{\pi}{4}ntb^2 \cot \frac{\phi}{2} \cosec^2 \frac{\phi}{2} d\phi.$$  

(158)

This gives the fraction of the total number of particles which are deviated between $\phi$ and $\phi + d\phi$. The fraction $\rho$ of the total number of particles which are deflected through an angle greater than $\phi$ is

$$\rho = \frac{\pi}{4}ntb^2 \cot \frac{\phi}{2}.$$  

(159)

The fraction $\rho$ which is deflected between the angles $\phi_1$ and $\phi_2$ is given by

$$\rho = \frac{\pi}{4}ntb^2 \left( \cot^2 \frac{\phi_1}{2} - \cot^2 \frac{\phi_2}{2} \right).$$  

(160)

Since in the experiment the number of scintillations appearing on a constant area of a zinc sulphide is counted, it is needed to express $dm$ given by Eq. (158) in another form for comparison with the experiment. Let $r$ be the distance from the point of incidence of $\alpha$-rays on the scattering material, and $Q$ the total number of $\alpha$-particles falling on the scattering material. The number $y$ of $\alpha$-particles falling on unit area which are deflected through an angle $\phi$ is then given by
\[ y = \frac{Qdm}{2\pi r^2 \sin \phi d\phi} = \frac{ntb^2 \cdot Q \cdot \csc^4 \phi/2}{16r^2}. \]  

Since \( b = 2NeE/mu^2 \), it is seen from this equation that the number of scintillations per unit area is proportional to the following quantities:

1. \( \csc^2 \phi/2 \), or \( 1/\phi^4 \) if \( \phi \) is small,
2. thickness of scattering material \( t \) provided this is small,
3. magnitude of central charge squared \( (Ne)^2 \),
4. inversely to \( (nu^2)^2 \), or to the forth power of the velocity.

In these calculations, it is assumed that the \( \alpha \)-particles scattered through a large angle suffer only one large deflection. This is valid provided the thickness of the scattering material is small. If, for example, the probability of a single deflection \( \phi \) in passing through a thickness \( t \) is 1/1000, the probability of two successive deflections each of value \( \phi \) is 1/10\(^6\), and is negligibly small.

Rutherford states that Geiger made a test of this theory of single scattering, finding that the distribution for particles deflected between 3\(^{\circ}\) and 150\(^{\circ}\) from a thin gold-foil was in substantial agreement with the theory.

Alternation of velocity in an atomic encounter — In the above consideration an \( \alpha \)-particle is assumed not to suffer an appreciable change of velocity in a large deflection. Rutherford treats now such a change. He supposes that only the two systems of \( \alpha \)-particle and atom are involved, and that the principle of conservation of energy-momentum applies, and also that there is no appreciable loss of energy or momentum by radiation.

Let \( u_1 = \) velocity of approach, \( u_2 = \) velocity of recession, \( M = \) mass of atom, \( V = \) velocity communicated to atom as a result of encounter, and \( \phi = \) angle between momenta \( mu_1 \) and \( mu_2 \). The conservation of momentum gives

\[ (MV)^2 = (mu_1)^2 + (mu_2)^2 - 2m^2u_1u_2 \cos \phi, \]

while that of energy

\[ MV^2 = mu_1^2 - mu_2^2. \]

One gets from these equations, by putting \( M/m = K \) and \( u_2 = \rho u_1 \), where \( \rho < 1 \),

\[ \rho = \frac{\cos \phi}{K + 1} + \frac{1}{K + 1} \sqrt{K^2 - \sin^2 \phi}. \]
In the case of an \( \alpha \)-particle deflected through an angle of \( 90^\circ \) by an encounter with an atom of gold of atomic weight 197, \( K = 49 \) and consequently \( \rho = 0.979 \), or the velocity of the \( \alpha \)-particle is reduced by 2\%. In the case of aluminum, \( K = 27/4 \) and \( \rho = 0.86 \) for \( \phi = 90^\circ \). Since the range of an \( \alpha \)-particle in air is approximately proportional to the cube of the velocity, it follows that an \( \alpha \)-particle of range 7 cm has its range reduced to 4.5 cm, after incurring a single deviation of \( 90^\circ \) in traversing an aluminum atom. Rutherford says that this is a magnitude to be easily detected experimentally. He mentions also that it is interesting to study the changes of velocity in the case of deflections due to a material of small atomic weight. The changes of velocity for large values of \( K \) are negligible.

Next, Rutherford considers the relative importance of single and compound scattering in determining the distribution of the scattered particles. Since the atom is supposed to consist of a central charge surrounded by a uniform distribution of charge of the opposite sign through a sphere of radius \( R \), the chance of encounters with the atom involving small deflections is very great compared with the chance of a single large deflection. The single scattering should therefore be compared with the compound scattering, which corresponds to what J. J. Thomson studied. This comparison will give also a comparison of the atomic model under consideration with that of Thomson.

According to Thomson's calculation, the average deflection \( \phi_1 \) due to the sphere of positive electricity (radius \( R \) and quantity \( Ne \)) is

\[
\phi_1 = \pi \frac{NeE}{4 \mu^2 R},
\]

and the average deflection due to the \( N \) electrons distributing uniformly throughout the atom is

\[
\phi_2 = \frac{16}{5} \frac{eE}{\mu^2 R} \sqrt{\frac{3N}{2}}.
\]

The mean deflection due to the both is taken to be

\[
(\phi_1^2 + \phi_2^2)^{1/2}.
\]

In a similar way, the average deflection due to an atom with nucleus can be calculated. In this case, the scattering is caused by the force \( X \) given above in the present subsection, and the deflection is given by
The Formation of Atomic Models

\[ \theta = \frac{b}{p} \left(1 - \frac{p^2}{R^2}\right)^{3/2}, \]

of which the average over the values of \( p \) is to be calculated. In the case of small \( p \), that is, in the case in which an particle passes near the center of the atom, \( \theta \) becomes very large. For \( \alpha \)- and \( \beta \)-particles the value of \( p \) for a large deflection is of the order of \( 10^{-11} \) cm. Because the chance of such an encounter is very small compared with the chance of small deflections, the average small deflection is practically unaltered if the large deflections are omitted. In this way the average small deflection is given by

\[ \phi_1 = \frac{3\pi b}{8R}. \]

This is three times the average deflection due to the corresponding sphere of positive electricity in an atom examined by Thomson. Combining the deflections due to the electric field and electrons, the average deflection becomes

\[ (\phi_1 + \phi_2)^{1/2} = \frac{b}{2R} \left(5.54 + \frac{15.4}{N}\right)^{1/2}. \]

Rutherford finds, later in his paper under consideration, that \( N \) is nearly proportional to the atomic weight, and is about 100 for gold. Needless to say, this value is too great. However, it is a value near the present-day value compared with the values assumed in his time because of the general supposition of \( N \) being proportional to the atomic weight. The effect due to the individual electrons expressed by the second term of the last equation is consequently small for heavy atoms compared with that due to the central charge. The average deflection per atom is then \( 3\pi b/8R \).

It is now possible to consider the relative effects on the distribution of particles due to single and to compound scattering. According to Thomson, the average deflection \( \theta_t \) after passing through a thickness \( t \) of matter is proportional to the square root of the number of encounters, and is given by

\[ \theta_t = \frac{3\pi b}{8R} \sqrt{\pi R^2 \cdot nt} = \frac{3\pi b}{8} \sqrt{\pi nt}, \]

where \( n \) is the number of atoms per unit volume. Since the probability \( p_1 \) for compound scattering that the deflection of the particle is greater than \( \phi \) is equal to \( e^{-\phi^2/\theta_t^2} \), one has
\[ \phi^2 = -\frac{9\pi^2}{64}b^2nt\log p_1. \]  

(162)

When single scattering alone is operative, the probability \( p_2 \) of a deflection greater than \( \phi \) is given by Eq. (159) as

\[ p_2 = \frac{\pi}{4}b^2nt\cot^2 \phi/2. \]  

(163)

The comparison between Eqs. (162) and (163) gives

\[ p_2 \log p_1 = -0.181\phi^2\cot^2 \phi/2. \]  

(164)

\( \phi \) is sufficiently small that \( \tan(\phi/2) \approx \phi/2 \), and therefore

\[ p_2 \log p_1 = -0.72. \]  

(165)

If \( p_2 = 0.5, p_1 = 0.24 \), and if \( p_2 = 0.1, p_1 = 0.0004 \). Namely, the probability for any given deflection is seen to be always greater for single than for compound scattering. The difference is especially marked when only a small fraction of the particles are scattered through any given angle. It follows hence that the distribution of particles due to encounters with the atoms is for small thicknesses mainly governed by single scattering.

Comparison of the theory with experiments — On the theory under consideration the value of the central charge \( Ne \) is an important constant, and it is desirable to determine this value for different atoms. This can be done by observing the small fraction of \( \alpha \)- or \( \beta \)-particles falling on a thin metal screen, which are scattered through angles between \( \phi \) and \( \phi + d\phi \). Rutherford discusses the following points:

(a) The "diffuse deflection" of \( \alpha \)-particles, i.e. the scattering of \( \alpha \)-particles through large angles (Geiger and Marsden).

(b) The variation of diffuse deflection with atomic weight of the radiator (Geiger and Marsden).

(c) The average scattering of a pencil of \( \alpha \)-rays transmitted through a thin metal plate (Geiger).

(d) The experiments of Crowther on the scattering of \( \beta \)-rays of different velocities by various metals.

(a) It was shown that about 1/8000 of the \( \alpha \)-particles from RaC falling on a thick plate of platinum are scattered back in the incident direction. This fraction is deduced on the assumption that the \( \alpha \)-particles are uniformly scattered
The Formation of Atomic Models

in all directions, the observations being made for a deflection of about 90°. Rutherford says that the form of experiments is not very suited for accurate calculation, but the scattering observed is about that to be expected if the atom of platinum has a central charge of about 100e.

(b) They measured the relative number of α-particles diffusely reflected from thick layers of different metals, under similar conditions. The numbers obtained are given in the table below, where $z$ represents the relative number of scattered particles, measured by the number of scintillations per minute on a zinc sulphide screen.

<table>
<thead>
<tr>
<th>Metal</th>
<th>Atomic weight $A$</th>
<th>$z$</th>
<th>$z/A^{3/2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead</td>
<td>207</td>
<td>62</td>
<td>208</td>
</tr>
<tr>
<td>Gold</td>
<td>197</td>
<td>67</td>
<td>242</td>
</tr>
<tr>
<td>Platinum</td>
<td>195</td>
<td>63</td>
<td>232</td>
</tr>
<tr>
<td>Tin</td>
<td>119</td>
<td>34</td>
<td>226</td>
</tr>
<tr>
<td>Silver</td>
<td>108</td>
<td>27</td>
<td>241</td>
</tr>
<tr>
<td>Copper</td>
<td>64</td>
<td>14.5</td>
<td>225</td>
</tr>
<tr>
<td>Iron</td>
<td>56</td>
<td>10.2</td>
<td>250</td>
</tr>
<tr>
<td>Aluminum</td>
<td>27</td>
<td>3.4</td>
<td>243</td>
</tr>
</tbody>
</table>

| Average     |                   |     | 233         |

On the theory of single scattering, the fraction of the total number of α-particles scattered through any given angle in passing through a thickness $t$ is proportional to $nA^2t$, if the central charge is assumed to be proportional to the atomic weight $A$. $t$ represents the greatest depth from which the scattered α-particles are able to emerge and affect the zinc sulphide screen. Since Bragg showed that the stopping power of an atom for an α-particle is proportional to $\sqrt{A}$, the value of $nt$ is proportional to $1/\sqrt{A}$. The number $z$ of α-particles scattered back from a thick layer is consequently proportional to $A^{3/2}$, or $z/A^{3/2}$ should be constant, in reasonably good agreement with the experiments shown in the table.

(c) Geiger made a careful determination of the scattering of α-particles passing through thin metal foils by the scintillation method, and deduced the most probable angle through which the α-particles are deflected, that is, the angle for which the number of scattered particles is maximum. A narrow pencil of homogeneous α-rays was used, and the total number of α-particles deflected through different angles was directly measured.
Examinations of Both Models and the Determination of Atomic Model

With Eq. (165), \( p_2 \log p_1 = -0.72 \), the probability \( q \) of the combined effects of single and compound scattering is given as \( q = (p_1^2 + p_2^2)^{1/2} \). If \( q = 0.5 \), it follows that \( p_1 = 0.2 \) and \( p_2 = 0.46 \). The probability \( p_2 \) of a single deflection greater than \( \phi \) is given by Eq. (163). Since in the experiments considered \( \phi \) is comparatively small, the following relation is obtained.

\[
\frac{\phi \sqrt{p_2}}{\sqrt{\pi nt}} = b = \frac{2N eE}{mu^2}.
\]

Geiger found that the most probable angle of scattering of the \( \alpha \)-rays in passing through a thickness of gold equivalent in stopping power to 0.76 cm of air was 1°40′. The angle \( \phi \) through which half the \( \alpha \)-particles are turned corresponds thus to 2° nearly. When \( p_2 = 0.46 \) is taken for the probability of single scattering, and the values

\[
t = 0.00017 \text{ cm}, \quad n = 6.07 \times 10^{22},
\]

\[
u(\text{average}) = 1.8 \times 10^9,
\]

\[
E/m = 1.5 \times 10^{14} \text{ e.s.u.}, \quad e = 4.65 \times 10^{-1},
\]

are substituted in the above relation, the value \( N = 97 \) comes out for gold.

For a thickness of gold equivalent in stopping power to 2.12 cm of air, Geiger found the most probable angle to be 3°40′. In this case \( t = 0.00047 \), \( \phi = 4°4′ \) and \( u(\text{average}) = 1.7 \times 10^9 \), and the value \( N = 114 \) comes out.

Geiger showed that the most probable angle of deflection for an atom was nearly proportional to its atomic weight. Rutherford deduces thus that \( N \) for different atoms should be proportional to their atomic weights, at any rate for atomic weights between gold and aluminum. Since the atomic weight of platinum is nearly equal to that of gold, it follows that from this considerations that the diffuse reflection through more than 90° from gold and the average small angle scattering in passing through gold foil are both explained by supposing about 100\text{e} for the central charge of the atom of gold.

(d) Experiments on scattering of \( \beta \)-rays — In Lenard’s experiments we have described before, the absorption of cathode rays was measured to deduce the size of Dynamiden. In Thomson’s work, the number of electrons in an atom was deduced, as we have described in §1 of the present section, from the spreading of a beam of incident rays on the assumption, that the scattering of incident particles was caused only by the electrons in the atom, made in
accordance with his model of the atom. Crowther measured the scattering of β-particles of different velocities by various materials, and tried to explain his results by Thomson's theory of compound scattering. Rutherford examines now how far the experimental results of Crowther can be explained on the general theory of single scattering.

On this theory, the fraction $p$ of β-particles turned through an angle greater than $\phi$ is given by Eq. (163). In Crowther's experiments $\phi$ is sufficiently small to put $\tan \phi/2 = \phi/2$. Consequently

$$\phi^2 = 2\pi nt b^2 \quad \text{if} \quad p = 1/2.$$  

On the theory of compound scattering, the chance $p_1$ that the deflection of the particles is greater than $\phi$ is given by Eq. (162) as

$$\frac{\phi^2}{\log p_1} = -\frac{9\pi^3}{64} nt b^2.$$  

In Crowther's experiments, the thickness of matter gave $p_1 = 1/2$, so that

$$\phi^2 = 0.96\pi nt b^2.$$  

The comparison of this with $\phi^2$ due to the theory of single scattering shows that they are identical in form but differ by a numerical constant. Rutherford points out that "it is thus clear that the main relations on the theory of compound scattering of Sir J. J. Thomson, which were verified experimentally by Crowther, hold equally well on the theory of single scattering". Rutherford indicates as an example that, if $t_m$ is the thickness for which half the particles are scattered through an angle $\phi$, Crowther showed that $\phi/\sqrt{t_m}$ and also $(mu^2)/E\sqrt{t_m}$ were constants for a given material when $\phi$ was fixed. These relations hold also on the theory of single scattering.

Rutherford states that not withstanding this similarity the two theories are fundamentally different in the distribution of scattered particles, which differs entirely on the two theories when the probability of deflection greater $\phi$ is small.

The distribution of scattered α-particles was found by Geiger's work mentioned above to be in substantial agreement with theory of single scattering, but cannot be explained on the theory of compound scattering. Since there is every reason to believe that the laws of scattering of α- and β-particles are very similar, the laws of their distribution should be the same as for small thicknesses of matter. On the other hand, the chance of large single deflections
Examinations of Both Models and the Determination of Atomic Model

219

for β-particles in passing through a given thickness of matter is much greater than for α-particles, because the value of \( \mu u^2 / E \) for the β-particles is much smaller than that for the α-particles. On the theory of single scattering, the fraction of the number of particles which are deflected through a given angle is proportional to \( t \). The number of particles which are undeflected is then proportional to \( 1 - kt \), \( k \) being a constant. As was shown by Thomson, on the theory of compound scattering the probability of deflection less than \( \phi \) is proportional to \( 1 - e^{-\mu t} \), where \( \mu \) is a constant for any given value of \( \phi \).

The correctness of this formula was tested by Crowther, by measuring electrically \( I/I_0 \) of the scattered β-particles which passed through a circular opening subtending an angle of 36° with the scattering material. If

\[
I/I_0 = 1 - e^{-\mu t},
\]

\( I \) should decrease very slowly at first with increase of \( t \). Crowther, using aluminium as scattering material, stated that the variation of \( I/I_0 \) was in good accord with this theory for small values of \( t \). However, if single scattering is present, as it undoubtedly is for α-rays, the curve showing the relation between \( I/I_0 \) and \( t \) should be nearly linear in the initial stages. Rutherford says that Marsden's experiments\(^{129} \) on scattering of β-particles certainly support such a conclusion of the theory of single scattering, too.

From the table given by Crowther of \( \phi/\sqrt{t_m} \) for different elements for β-rays of velocity \( 2.68 \times 10^{10} \) cm per sec, Rutherford calculates on the theory of single scattering the values of the central charge \( Ne \). It is supposed, as in the case of the α-rays, that for the given value of \( \phi/\sqrt{t_m} \) the fraction of the β-particles deflected by single scattering through an angle greater than \( \phi \) is 0.46 instead of 0.5. The calculated values of \( N \) are given in the following table.

<table>
<thead>
<tr>
<th>Element</th>
<th>Atomic weight</th>
<th>( \phi/\sqrt{t_m} )</th>
<th>( N )</th>
<th>Present-day ( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium</td>
<td>27</td>
<td>4.25</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>Copper</td>
<td>63.2</td>
<td>10.0</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>Silver</td>
<td>108</td>
<td>15.4</td>
<td>78</td>
<td>47</td>
</tr>
<tr>
<td>Platinum</td>
<td>194</td>
<td>29.0</td>
<td>138</td>
<td>78</td>
</tr>
</tbody>
</table>

\(^{129}\)E. Marsden, Phil. Mag. 13 (1909), 909.
The values of $N$ for gold deduced from scattering of the $\alpha$-rays are in two calculations 97 and 114. These numbers are somewhat smaller than the value $N = 138$ given above for platinum, whose atomic weight is not very different from gold. Taking into account the uncertainties involved in the calculation from the experimental data, Rutherford concludes that "the agreement is sufficiently close to indicate that the same general laws of scattering hold for the $\alpha$- and $\beta$-particles, notwithstanding the wide differences in the relative velocity and mass of these particles". He points out also that, as in the case of $\alpha$-rays, the value of $N$ should be most simply determined by measuring $\beta$-particles scattered through a large angle.

Still more, he mentions that Schmidt's experimental deductions\(^{130}\) of the absorption of $\beta$-rays by various elements are in agreement with the theory of single scattering.

In this way, Rutherford makes it clear that the large single deflections of $\alpha$- and $\beta$-particles are mainly due to their passage through the strong central field. He reasons that, because of the mass, momentum and kinetic energy of the $\alpha$-particle are very large compared with those of an electron in rapid motion, it does not seem possible from dynamical considerations that an $\alpha$-particle can be deflected through a large angle by a close approach to an electron, even if the latter is in rapid motion and constrained by strong electrical forces.

Rutherford takes then the problem of the question of the extent of the distribution of the central charge. If, in a collision, the central charge $Ne$ is supposed not to act as a whole but to act as $N$ constituent units, the fraction of the $\alpha$-particles scattered through a given angle is proportional to $Ne^2$ instead of $(Ne)^2$. In this consideration, the influence of mass of the constituents is neglected. Now, since the central charge for gold is found to be 100 if it acts as a whole, it should be 10,000 if it is supposed to act singly. Then, the mass of the constituent particle would be small compared with that of the $\alpha$-particle, and the production of large single deflections would not occur at all, the effect of compound scattering becoming much more important than that of single scattering. Rutherford gives thus the conclusion that the large single deflections are due to the central charge as a whole. At the same time, he mentions that the experimental evidence is not precise enough to negative the possibility that a small fraction of the positive charge may be carried by its

\(^{130}\)H. W. Schmidt, Ann. d. Phys. 23 (1907), 671.
Examinations of Both Models and the Determination of Atomic Model

satellites. He says that evidence on this point could be obtained by examining the difference between large single deflections of $\alpha$- and $\beta$-rays.

Rutherford points out that, though the general data available indicate that the value of the central charge is proportional to the atomic weight for elements heavier than aluminium, it will be of great interest to examine whether such a simple relation holds also for the lightest elements, and that in those cases the general theory of single scattering will require modification because the mass of the deflecting atom is not very different from that of the $\alpha$-particle.

Rutherford states then about Nagaoka's model of the atom as follows — "It is of interest to note that Nagaoka has mathematically considered the properties of a ‘Saturnian’ atom which he supposed to consist of a central attracting mass surrounded by rings of rotating electrons. He showed that such a system was stable if the attractive force was large. From the point of view considered in this paper, the chance of large deflection would practically be unaltered, whether the atom is considered to be a disk or a sphere”.

On the central body Rutherford remarks an idea in anticipation, which may be said to be somewhat similar to the present-day knowledge of the structure of atomic nuclei. In other words, he says that the approximate value 100e found for the central charge of the atom of gold with the atomic weight 197, is about that to be expected if the atom of gold consisted of 49 atoms of helium, each carrying a charge 2e, and that this may be only a coincidence, but it is certainly suggestive in view of the expulsion of $\alpha$-particles from radioactive matter.

Rutherford is very careful about the sign of the central charge. It is noteworthy that he makes no such inference at all as to say the positiveness of the central charge by reason of the neutralization brought about by the surrounding negative electrons. He remarks that the question of sign may be settled by the reduction in the velocity of the $\beta$-particle, which should be far more marked with a positive than with a negative center, and that the high velocity of expulsion of $\alpha$-particles may be accounted for by the repulsive force due to the central charge of positive electricity.

Rutherford’s theory we have described above brings thus the complete determination to the question of the models of the atom.

§5. *About the number of electrons in an atom-II*

On account of the final determination by Rutherford of the atomic model, the problem of the number of electrons in an atom came to be examined again.
from a new viewpoint. It became necessary to reexamine Thomson’s theory described in §1, because that theory had been set up on the basis of the model of the atom without nucleus. In such an examination, the scattering and dispersion of X-rays by an atom would bring almost no direct problem, but those of charged particles passing through atoms should be treated separately for the cases of those due to the central charges and of those due to the electrons. Those due to the electrons cause almost no deflection, but certain loss of energy, called the ionization loss, on account of which charged particles passing through matter stop within some range.

The number of electrons in an atom can be determined by means of the correct theory of ionization loss. Although the theory can be arrived at of course with the use of quantum mechanics, Bohr\textsuperscript{131}) obtained within the framework of the classical theory the number of the electrons in an atom of light elements from the loss of velocity of charged particles.

It should be noted that this work was done just before Bohr’s epoch-making work of applying the quantum theory to atoms. Particularly in the paper under consideration the hydrogen atom is shown to have only one electron. This finding brought the decisive difficulty of the stability and permanence of atoms to the classical theoretical treatment of the model of the atom with nucleus. Bohr’s theory of the atom based fundamentally on the determination of the atomic model of this kind.

In 1912 Whiddington\textsuperscript{132}) examined, following Thomson’s theory, the decrease of velocity of charged particles, and showed that Thomson’s theory was in good agreement with experiments. In Thomson’s theory α-particles were supposed to be of great dimension, and accordingly to collide with each atom as a whole. Rutherford’s experiment made clear that the α-particle was identical with the nucleus of helium atom and of very small dimension, being different from the β-particle only in charge and mass. According to this conception, an elaborate theory of the absorption and scattering of α-rays was set forth by C. G. Darwin.\textsuperscript{133})

Darwin supposed the Coulomb force for the interaction between an α-particle and each electron in an atom, and made use of some assumptions to obtain results for the scattering and absorption in good agreement with experiments. The assumptions used were however open to objections. In his

\textsuperscript{131})N. Bohr, Phil. Mag. 23 (1913), 10.
\textsuperscript{133})C. G. Darwin, Phil. Mag. 23 (1912), 907.
theory, the loss of velocity of a charged particle was assumed to be due to a transfer of kinetic energy to the electrons in the atoms with which it collides. The forces keeping the electrons in an atom were assumed to be negligible during the short collision time. The motion of each electron during the collision could thus be calculated, to obtain the loss of energy of the $\alpha$-particle. However, if the total loss of energy due to all the electrons in the matter was integrated, there would result an infinitely great value.

In the case of Thomson's theory, this infinity was avoided by cutting off the action of each electron at a distance comparable with the distances among the electrons in an atom. This procedure was taken on the consideration that at such a distance the actions due to many elections would result in null effect. The influence of these simultaneous actions is, however, ineffective for the loss of velocity, though effective for the deflection, of a charged particle. This is because the force exerted by a charged particle of high velocity on an electron, and consequently the energy transferred to the latter by the collision is irrespective of the actions due to the other electrons.

Darwin avoided this difficulty with another method. In other words, he assumed that the atomic binding force did not act on any electron during the collision of short time, and that the velocity of the $\alpha$-particle was not altered if its path did not across the atom. On these assumptions Darwin found from the comparison of his result with experiments that the diameter of atoms decreased with increasing atomic weight, being several times greater for the lightest atoms, and several times smaller for the heaviest elements, than the values generally adopted.

Now, in the paper under consideration, Bohr criticizes that it is not justifiable to take the surface of an atom as the limit of the effect of the electrons in the atom. In other words, he points out that, though outside an atom the positive charge and negative ones in the atom can be considered to result in a neutral effect, the decrease of velocity of a particle of high velocity depends only on the motion of the individual electrons during the collision, and not on the total force exerted on the particle by the whole atom, the latter force producing only the scattering of the particle.

It is very interesting that Bohr bases his theory fundamentally on the vibration of atomic electrons. We see in it an expectation of quantum theoretical atomism. Bohr introduces the following idea of a natural limit for the effect of the electrons on the velocity of the moving particle. The electrons will have a sort of vibratory motion, if they are disturbed by an impulse from outside.
The motion of the electrons will be materially altered during the collision, and consequently the particle will lose its energy, if the time of vibration of the electrons is of the same order of magnitude as the time of collision, i.e., the time which the particle takes to travel a distance of the same order of magnitude as the shortest distance $p$ of the electron from the path of the particle.\cite{134} The energy given by the particle to the electrons will decrease very rapidly, if the distance of the particle from the electrons is so great that the time of collision is great compared with the time of vibration. The effective limit will thus depend purely on the frequency of the electrons and the velocity of the particle. The limit in question may, for the same velocity of the particle, be very different for individual atomic electrons. The limit will be very small for some of the electrons in elements of high atomic weight because of their very high frequencies, so that such elements will show comparatively small absorption.

Bohr points out that the theory of the decrease of velocity of moving charged particles on passing through matter bears a great analogy to the ordinary electromagnetic theory of dispersion. In fact, making use of the information about the number and frequency of atomic electrons obtained from the theory of dispersion, Bohr calculates the absorption of $\alpha$-rays for the lightest elements in very close agreement with experiments. Since the decrease in the effect of the electrons with increasing frequency is much more rapid for the dispersion than for the loss of velocity of moving particles, the latter is more suitable than the former for getting information about atomic electrons of high frequency.

Let $M$ and $E$ be the mass and charge of a particle, respectively, and let us assume that the particle has a velocity $V$ before the collision. The collision with a free electron is treated in the same way as that in Thomson’s theory described in §1. Let $p$ be the distance from the electron to the path of the particle before the collision, and $2\theta$ be the angle through which the direction of the relative motion is deflected by the collision. Assuming that the electron is initially at rest, one has

$$\sin^2 \vartheta = \frac{1}{1 + \frac{p^2}{\lambda^2}},$$

(166)

where

$$\lambda = \frac{eE(M + m)}{V^2 m M}.$$ 

(167)

\cite{134} J. J. Thomson, Phil. Mag. 23 (1912), 454.
The energy transferred to the electron is

\[ Q_0 = \frac{2mM^2V^2}{(m + M)^2} \sin^2 \theta. \]  

(168)

The displacement of the electron at the moment of the nearest approach of the particle to the electron is \((eE/mV^2) \cos \theta\). If \(p \gg \lambda\), the velocity of the electron after the collision is almost perpendicular to the path of the particle. Now, as regards an atomic electron, if the frequency of the electron is small, the time of vibration is very long compared with the time of collision for collisions of \(p \sim \lambda\). This corresponds to the case of the lightest elements. Consequently, in this case collisions of \(p \gg \lambda\) only should be considered, and the calculation is simplified.

In the figure, AB represents the path of the particle, which is very nearly a straight line. A is the position of the particle at the time \(t\), and C is the mean position of the electron. BC \(\perp\) AB, \(BC = p\), B being the position of the particle at the time zero, so that \(AB = Vt\). The force exerted on the electron in the direction CB is given by

\[ F_1 = eE \frac{BC}{AC^3} = \frac{eEp}{(V^2t^2 + p^2)^{3/2}} m \cdot \phi(t). \]

The equation of the motion of the electron perpendicular to the path of the particle is
\[
\frac{d^2 x}{dt^2} + n^2 x = \phi(x),
\]
where \( n \) is the frequency corresponding to the binding force due to the atom. The energy transferred to the electron by this force during the collision is

\[
Q_1 = \frac{m}{2} \left( \frac{dx}{dt} \right)^2 + \frac{mn^2}{2}x^2,
\]
where \( dx/dt \) and \( x \) can be expressed in terms of integrals over time \(-\infty \to t\).

The force exerted on the electron in the direction parallel to the path of the particle is given by

\[
F_2 = eE \frac{AB}{AC^2} = \frac{eEv}{(V^2t^2 + p^2)^{3/2}} = m\psi(t).
\]

Let \( Q_2 \) be the energy transferred to the electron by this force during the collision. The total energy transferred to the electron is

\[
Q = Q_1 + Q_2 = \frac{2e^2E^2}{mV^2p^2}P\left(\frac{np}{V}\right),
\]
where \( P(x) \) is a function which has the values \( P(0) = 1 \) and \( P'(0) = 0 \) and decreases rapidly for increasing values of \( x \).

Let us assume that the number of atoms per unit volume is \( N \), and that each atom contains \( r \) electrons of frequency \( n \). Let \( a \) be a constant such that \( \lambda \ll a \ll V/n \). The total energy \( dT \) transferred to the electrons when the particle travels a distance \( dx \) is

\[
dT = Nr \left[ \int_0^a Q_0 2\pi pdp + \int_a^\infty Q 2\pi pdp \right] dx.
\]

Inserting the expressions of \( Q_0 \) and \( Q \) given above one gets

\[
dT = \frac{4\pi e^2E^2Nr}{mV^2} \log \left( \frac{V^3kMm}{neE(M + m)} \right) dx,
\]
where

\[
\int_0^\infty \log zP'(z)dz = -\log k.
\]

\( k \) is calculated to have the value

\[
k = 1.123.
\]
Examinations of Both Models and the Determination of Atomic Model

If each atom contains \( r \) electrons of frequencies \( n_1, n_2, \ldots, n_r \), one has

\[
dT = \frac{4\pi e^2 E^2 N}{mV^2} dx \sum_{s=1}^{r} \log \left( \frac{V^2 kMm}{n_e E(M + m)} \right) .
\]

Since \( dT \) is equal to the decrease in the kinetic energy of the particle, \((1/2)MV^2\), one gets

\[
\frac{dV}{dx} = -\frac{4\pi e^2 E^2 N}{mMV^2} dx \sum_{s=1}^{r} \log \left( \frac{V^3 kMm}{n_e E(M + m)} \right) .
\]

In the derivation of Eq. (173), only the interaction of the electrons with the central charge of the atom and the particle is considered. For large values of \( V \) one obtains the relation of the same form as that of J. J. Thomson and Whiddington. In terms of \( V_0 \), the velocity of the particle at \( x = 0 \), and \( a = -4(V^3dV/dx)_0 \), it is given by

\[
V_0^4 - V_x^4 = ax .
\]

If the number of electrons in an atom is proportional to the atomic weight, and electrons of higher frequencies are contained in an atom of heavier atomic weight, Eq. (173) accounts for the following facts well: the absorption of \( \alpha \)-rays for equal weight of matter per cm\(^2\) decreases with increasing atomic weight, and the relative absorption of \( \alpha \)-rays of high velocity is relatively great for elements of large atomic weight.

Bohr compares his result first with experiment of \( \alpha \)-rays. On account of small scattering, \( \alpha \)-rays are more convenient for the study of absorption than \( \beta \)-rays. He cites the experimental relation found by H. Geiger

\[
V^3 = KR ,
\]

where \( R \) is the range of \( \alpha \)-rays in air and \( K = 1.10 \times 10^{27} \), and says that this relation is very accurate. Bohr makes use of the measurements made by T. S. Taylor of the relative absorption coefficients of various elements for \( \alpha \)-rays of various velocities. They are given in the following table.

---

According to Eq. (174) it is known that the ranges 2.24 and 4.87 correspond to the velocities $1.35 \times 10^9$ and $1.75 \times 10^9$, respectively. From this Bohr deduces the values of $-dV/dx$ given in the following table.

<table>
<thead>
<tr>
<th>Range in air</th>
<th>2.24</th>
<th>4.87</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen</td>
<td>0.267</td>
<td>0.224</td>
</tr>
<tr>
<td>Air</td>
<td>1.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Aluminium</td>
<td>$1.6 \times 10^3$</td>
<td>$1.75 \times 10^3$</td>
</tr>
<tr>
<td>Tin</td>
<td>$2.33 \times 10^3$</td>
<td>$2.56 \times 10^3$</td>
</tr>
<tr>
<td>Gold</td>
<td>$4.71 \times 10^3$</td>
<td>$5.57 \times 10^3$</td>
</tr>
<tr>
<td>Lead</td>
<td>$3.06 \times 10^3$</td>
<td>$3.53 \times 10^3$</td>
</tr>
</tbody>
</table>

$Hydrogen$ — Into Eq. (173) Bohr inserts the values $e = 4.65 \times 10^{-10}$, $E = 2e$, $e/m = 5.31 \times 10^{17}$, $E/m = 1.46 \times 10^{14}$ and $N = 2.59 \times 10^{19}$ (at 76 cm and 20°C). He uses then the result deduced by C. and M. Cuthbertson from experiments on the refraction and dispersion in hydrogen according to Drude’s theory.\(^{137}\) They found that a hydrogen molecule contains 2 electrons, each being of frequency $n = 2.21 \times 10^{16}$. This means that $r = 2$ and $n_1 = n_2 = 2.21 \times 10^6$, with the use of which Bohr obtains

$$
\frac{dV}{dx} = -4.9 \times 10^7 \quad \text{for} \quad V = 1.351 \times 10^9,
$$

$$
\frac{dV}{dx} = -2.6 \times 10^7 \quad \text{for} \quad V = 1.75 \times 10^9.
$$

\(^{137}\)C. and M. Cuthbertson, Proc. Roy. Soc. A\textbf{83} (1909), 93. It should be noted that P. Drude, Ann. d. Phys. \textbf{14} (1904), 714, obtained in a similar way a value somewhat smaller than 2 for the number of electrons in a hydrogen molecule, as we have already mentioned in §1.
These values are in close agreement with the values of \( \frac{dV}{dx} \) in the above table \(-5.4 \times 10^7\) and \(-2.7 \times 10^7\), respectively. Namely, they are in good agreement with the result deduced by means of optics. If the hydrogen molecule is assumed to contain more than two electrons, the extra electrons must have extremely high frequencies of \( n \gtrsim 10^{18}\), in order for the absorption not to amount to more than about 10\%. Such a value of \( n \) seems difficult to be reconciled with experiments on characteristic X-rays. Bohr confirms thus that according to Rutherford’s atomic model the hydrogen atom can have only one electron. This has an epoch-making importance in atomic physics, as we have stated before.

**Helium** — The only experiment on this gas was that of Adams\(^{138}\) with \( \alpha \)-rays from polonium. With the use of the result \( r = 2 \) (the value directly obtained being 2.3) deduced by Cuthbertson from optics, Bohr calculates \( \frac{dV}{dx} \) from Eq. (173) and gets values about 10\% smaller than the experimental ones. Considering that this disagreement is due to a little contamination and the neglected corrections in calculation, Bohr concludes that the helium atom contains two electrons.

**Oxygen** — For the ratio of the absorption in oxygen to that in air Adams found the value 1.03. Bohr deduced thus

\[
\frac{dV}{dx} = -2.07 \times 10^8 \quad \text{for } V = 1.35 \times 10^9 ,
\]

\[
\frac{dV}{dx} = -1.24 \times 10^8 \quad \text{for } V = 1.75 \times 10^9 .
\]

With the use of Eq. (173) he then gets for the respective cases

\[
\sum_{s=1}^{r} (\log(n_s \times 10^{-19}) + 0.59) = -47 ,
\]

\[
\sum_{s=1}^{r} (\log(n_s \times 10^{-19}) - 0.18) = -61 .
\]

From these equations he obtains by subtraction \(0.77r = 14\), or \(r = 18\). Bohr says that according to Rutherford’s theory an oxygen molecule should contain 16 electrons in satisfactory agreement with the above value of \( r \).

\(^{138}\)E. P. Adams, Phys. Rev. 24 (1907), 113.
From experiments on dispersion of light 4 electrons in an oxygen molecule are known to be of frequency $2.15 \times 10^{16}$. Inserting this value into the equations given just above, Bohr gets

$$\sum_{s=5}^{r} \log(n_s \times 10^{-19}) = -34.$$  

Assuming that the other 12 electrons are of one and the same frequency $n_s = n'$, he obtains $n' = 0.61 \times 10^{18}$. From the fact that the velocity of an electron to excite the characteristic X-rays of an element is given by $A \times 10^8$ cm/s, $A$ being the atomic weight, Bohr deduces $n = 1.1 \times 10^8$ for the characteristic X-rays of oxygen by applying Planck's quantum theory, and concludes that this value is in remarkable agreement with that of $n'$ obtained above. The application made here by Bohr of the quantum theory is a very interesting one, and we shall treat it in Vol. II. \[139\]

**Elements of high atomic weight** — In a similar way to the case of oxygen, Bohr obtains the following values of $r$ for aluminium, tin, gold and lead.

<table>
<thead>
<tr>
<th>Substance</th>
<th>$r$</th>
<th>Atomic weight $A$</th>
<th>Today's $r$</th>
<th>$A/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium</td>
<td>14</td>
<td>27</td>
<td>13</td>
<td>13.5</td>
</tr>
<tr>
<td>Tin</td>
<td>38</td>
<td>119</td>
<td>50</td>
<td>59.5</td>
</tr>
<tr>
<td>Gold</td>
<td>61</td>
<td>197</td>
<td>79</td>
<td>98.5</td>
</tr>
<tr>
<td>Lead</td>
<td>65</td>
<td>207</td>
<td>82</td>
<td>103.5</td>
</tr>
</tbody>
</table>

According to Rutherford's theory $r$ is expected to be equal to about half the atomic weight. About the discrepancy from this expectation, Bohr says that the neglected corrections in the calculation may increase with atomic weight.

About the absorption of electrons of high velocity — Bohr compares the experimental results obtained by R. Whiddington\[140\] of the decrease of velocity of cathode rays in aluminium, tin, gold and lead, with the results calculated from Eq. (173) with the use of the values of $r$ and $\sum \log n_s$ deduced from the absorption of $\alpha$-rays. The observed and calculated results agree in the orders of magnitude, but differ greatly in the values, the calculated values being about three times larger than the experimental values. Bohr says that the origin of this discrepancy seems difficult to be accounted for.

\[139\]Translator's note: Cf. Sec. 4, §1, Chap. 2, Vol. II.
Examinations of Both Models and the Determination of Atomic Model

About the experimental results obtained by W. Wilson\(^{141}\) for very hard \(\beta\)-rays, Bohr obtains a better agreement.

From the present-day viewpoint — The expectation from Rutherford's theory of the number of electrons in an atom \(r = A/2\) is of course not correct. \(r\) means the present-day atomic number. Rutherford's \(r\) is too smaller, but gives fairly good values. The true \(r\) is given its foundation by Bohr's quantum theory\(^{142}\) in 1913, and is determined experimentally by Moseley\(^{143}\) in the same year.

Concluding Remarks

In the history of scientific cognition from ancient times, there are two problems which are most important, most interesting and most characteristic. One is the solar system, and the other is the structure of atoms. Both the two do not show their respective structures directly to our intuition. On account of this human beings could not but pass along very winding courses to arrive at the cognition of these structures. Moreover, it is also important that the cognition of them established the core of the foundation of the present-day physics. In other words, that of the former gave birth to the Newton mechanics and that of the latter to quantum mechanics.

True studies of the development of science make it clear, that the process of cognition by human beings never goes on in such a way of knowing mathematical laws through the systematization of given sensuous experience, as Machist or Kantian idealists and other idealists say.

As has been set forth in my book "Problems in Dialectic", science does not end in phenomenological descriptions of this kind, or in other words, scientific cognition does not end in a phenomenological stage of describing sensuous experience, but goes on, through the intermediation of a substantialistic stage of knowing what the object consists of or what structure it has, finally to an essentialistic stage of grasping universal and necessary cognition.

In the process of the formation of the Newton mechanics, the astronomy in ancient Egypt and Babylonia described the movement of each planet on the dome of the sky from the viewpoint of the everyday intuition on the earth,

---


\(^{142}\)N. Bohr, Phil. Mag. \textbf{26} (1913), 1, 476, 857.

\(^{143}\)H. G. J. Moseley, Phil. Mag. \textbf{26} (1913), 1024; \textbf{27} (1914), 703.
arriving at the mathematical law of its periodicity, but science did not end with this. In the time of ancient Greece, scientific thinking fled out of the earth into the structure of the universe, or the structure of the system of the planets, giving at last the fruit of the Kepler theory. This stage is a substantialistic stage. Further, by means of the cognition of the mutual interaction between material bodies, scientific thinking arrived at the Newton mechanics as an essentialistic stage.

The formation of quantum mechanics was along a similar way. Quantum mechanics was not formed by means of systematizing the phenomenological knowledge, such as the law of energy distribution of thermal radiation and that of spectral lines, as mere description of experience, but was formed after the establishment of the cognition of the substantialistic photon, electron and atomic structure. The determination of the atomic model makes an important stage in the history of physics. Being said in relation to nuclear physics, this stage corresponds to the ten odd years from 1930 to the present time.

Seen from the viewpoint of the development of the relativity theory, the formation of quantum mechanics is also an important problem. The substantialistic idea of the elastic ether was presented in the beginning of the 19th century. Though it brought forth various contradictions, they were solved by the establishment of the more clear substances, the atomic structure on one hand and the quantum of light on the other hand, and the essentialistic stage of the relativity theory was arrived at finally.

In Chap. 1 of the present volume, descriptions have been given of the progress of making it clear that phenomena of light should not be considered elastic vibrations of the ether, but should be understood by means of the substantialistic idea of the emission of particles of light according to the quantum theory. Volume II will treat the intensification of the contradiction between the concepts of wave and particle, which gave birth to quantum mechanics.

In Chap. 2 of the present volume, the progress of the gradual elucidation of the structure of atoms has been described. It has a very important epistemological meaning, that in this stage the substantial structure of atoms was determined clearly to a certain extent without any use of quantum theoretical idea. This is by virtue of the fact, that the result of quantum mechanical calculation of the normal Zeeman effect or the Rutherford scattering formula does not contain the Planck constant \( h \), and hence agrees completely with the result of classical theoretical calculation of it. It is noted also that, for other problems such as the question of the number of electrons in an atom,
classical mechanics does not give very wrong results in the meaning of approximation. This is related to the correspondence principle, and is connected with the reality of quantum mechanical objects, showing the definite significance of substantialistic stage. One sees clearly that a popular view taken by philosophers, which doubts if the object of quantum mechanics is real, is nothing but a great misunderstanding.

The reality of the atom has quite the same meaning as that of the solar system. The latter was doubted very much around the time of Galilei. The reality of things which give no direct sensation cannot be believed, as long as the standpoint of the Mach–Kantian idealism is taken. The naive realism cannot overcome such an idealism.

Mr. Takasi Ide, a sincere philosopher, confessed honestly his doubt about this problem, saying “Although I assert the world of the physics after Newton, and believe of course without doubt the existence or falling of a stone as an objective fact irrespective of my consciousness, I cannot understand what is meant by the reality of elements or atoms which are said to constitute it.” He described also as follows — “According to phenomenologists, even if one speaks of atoms he cannot make any sensuous experiment on them, but rather treats them only in his Gedankenexperiment. Hence, atoms are not so certain as stones we can grasp. The only thing we can understand may be the phenomenon or law of the motion of atoms. I think it possible to speak of the phenomenon or law of their motion. It could not be said, however, that ‘there is an atom’ in the same meaning as ‘there is a stone’. Thus, it seems to me that we go back again to Kant. I wonder whether there would be a firm basis for rejecting this Kantism as an idealism from the viewpoint of materialism, since there is a great difference, I think, in the standpoint or in the manner of presenting the problem . . . .”

It means a naive realism to believe merely objective existence outside consciousness without intermediation, and cannot be proof against Kant’s critique. In this way of thinking the existence even of a stone cannot be asserted academically.

The Kantian epistemology lacks the element of practice in cognition. In this meaning the theory of technology is the very epistemology of natural science (cf. “Problems in Dialectic”). The viewpoint of success and failure makes the intermediation between consciousness and nature. It is the dialectic materialism that bases epistemology upon practice. In other words, our action
will succeed as long as our cognition of the object is correct, while it will fail
as long as our cognition of the object is wrong.

Reality can never be obtained by the systematization of given experience. If we practice various actions upon an object with an image of it and get the expected result, the image has reality. The process of cognition must therefore be placed at the center of epistemology. In whatever manner a recognized result is merely dealt with, there will come out only arbitrary interpretations based on the policy of no success and no failure, and its reality will be lost. To the contrary, such a result strange to the everyday common idea as quantum mechanics or the cognition some ages ago of the revolving earth, can best be guaranteed on its reality by being based on practice and by being examined for the process of its formation through a series of success and failure. The existence of a stone and that of an atom can thus be said to have quite the same reality. Rocket V-2 took a photograph of the round earth, an electron microscope imaged some molecules, and nuclear physics gave rise to atomic bombs, just opening the so called atomic age.

Explanatory Notes

(Added to the republished edition in 1972)

Masayuki Nagasaki

(1) Mituo Taketani's book "The Formation and Logic of Quantum Mechanics Vol. I — The Formation of the Atomic Model — " was published in Jun., 1948 by the Ginza Pub. Co. It went out of print, however, because of the failure soon after of the company. The republication of this book was greatly desired by many persons, but has not been carried out for a long time. It is a great pleasure to have now its republished edition after the revision of some misprints, in spite of the use of photocomposition.

The purpose of this book consists, as is stated in the preface, in doing "the work of tracing the correct process of formation of the atomic theory, and thereby drawing its logic". This purpose is to be accomplished first in order to "make clear the logical structure" of atomic physics as "one of the greatest achievements of human beings", and thereby "put the outcome in the property of thought". This is because the great experience in the process of
development of atomic physics could not make a guiding principle in future practice of human beings, if it is not grasped as to its logic.

The present volume was not Taketani's first work on the history of science. In his paper “One Stage of Modern Physics — On the academic achievements of Mr. and Mrs. Joliot Curie, Nobel Prize Winners”, published in the Apr. and May issues (1936) of “Sekai Bunka” (World Culture), Taketani traced the important points of the process, starting from the view of matter in the 19th century and coming, through molecular and atomic theories, studies of radioactivity, atomic physics, quantum mechanics and nuclear physics, to the outset of the theory of elementary particles due to the presentation of the Yukawa meson theory, and deduced thereby an “epistemological significance”.

In this paper Taketani presented a full account of his famous theory of three stages, that the development of science takes place through the three logical stages: phenomenological, substantialistic and essentialistic stages. This paper describes first that, in the formation of Newton mechanics, the lawlessness in the phase of Tycho Brahe's time got a certain lawfulness by means of the introduction of the substantialistic model of the solar system, and rose to a phenomenological law called Kepler's law, in which the element of “human action, experiment, labour” was brought by Galilei’s physics born of the development of productive technology in the Renaissance, to make Newton introduce the essential concept of the force, who arrived finally at the law given by the fundamental equation of Force = Mass × Acceleration. The paper under consideration describes then, about “atoms and molecules as the model of matter” the process of “chemistry before the introduction of them (a phenomenological stage), the introduction of chemical atoms and molecules, the determinate meaning of Brownian motion, and the passage from phenomenological thermodynamics to kinetic theory of molecules”, about “the structure of atoms and molecules” the process of “nonunderstandable spectral lines (a phenomenological stage), Rutherford’s atomic model, Bohr’s solution still in phenomenological and post hoc meaning, and the establishment of quantum mechanics by means of the reconsideration of it from the viewpoint of physical quantity”, and about “nuclear physics” the process of “the chaos dominating up to 1930, and the finding in the five years since then of such elementary particles as neutron, positive electron, (neutrino), as the substantial elements to construct models, that opened the way to its second phase”.

This paper was included in Taketani’s book “Problems in Dialectic” published in 1946 [transcribed together with “Problems in Dialectic, 2nd series” in
“Collected Works of M. Taketani, Vol. 1” (Keiso Pub. Co., 1968), which will be abbreviated to “Collected Works 1” hereafter. Taketani says in his paper in 1946 “Modern Physics and Epistemology” [“Collected Works 1”, p. 27] that his indication of the importance of the substantialistic stage in the development of scientific cognition was made already in 1934, as follows — “In my graduation thesis in 1934, I pointed out that the cognition in physics must pass, in its development, an important and noticeable stage, that is, the stage of questioning what there is and what structure it has, and must make use of this stage as a stepping stone, and that the state of nuclear physics at that time was just in such a stage, and the contradictions were to be absorbed along this line.”

In Aug. 1948 Taketani published the paper “On the Formation of Newton Mechanics” in the journal “Kagaku” (Science) [“Collected Works 1”, p. 80]. In this paper he developed more thoroughly the arguments given in the paper “One Stage of Modern Physics”. This work was done with his cognizance of the problem which he described in the paper under consideration as follows — “We are confronted, in developing the meson theory which is one of the central problems in modern physics, with the serious situation whether quantum mechanics holds good here or not. Some powerful reconsideration and policy are wanted. It is said that one cannot be always lucky. However, doesn’t the process of development of science in the past give us anything?”

In this paper, Taketani made clear that “science progresses by being governed by the three of technology, constitution of nature itself and mode of thinking”, and thereby presented the right distinguishing features in studying the history of science, establishing the methodology for the history of science. Also in this paper, Taketani pointed out the existence of “the three forms of the movement from substantialistic to essentialistic stage”, and deduced at the same time his well-known conclusion, that “nuclear physics has been in a substantialistic stage of introducing substances, and the present difficulty can be said to lie in seeking for, by making substantialistic systematization first by virtue of the development of the meson theory, ways to an essentialistic stage”. As to a very great contribution of this analysis to the development of the theory of elementary particles in Japan, the reader is referred to S. Sakata’s paper “The Modernity of Modern Science” [included in “Iwanami Series of Lectures on Philosophy 6” (1968)]. The papers cited above of Taketani and Sakata, except “Modern Physics and Epistemology”, were edited systematically, on being translated into English, in Supplement to Progress of Theoretical Physics,
Examinations of Both Models and the Determination of Atomic Model

No. 50 (1971) with the title "Philosophical and Methodological Problems in Physics", to which the reader is also cordially referred.

This book was written with those things mentioned above for a background. For this reason Taketani says in Preface to the present volume "I have described in detail the method of our study in my book "Problems in Dialectic", and I will be glad if the reader refers to it." As a supplement, Taketani's following words may be quoted, which were in answer to a question asked by the present writer [Sep., 1972] — "At that time there was little work of this kind, other than Amano's one. Though Amano's work was well done, it was lacking in such points. This was so also throughout the world. Moreover, these points are the most important ones seen from my philosophical viewpoint. In other words, they are the most important points in making clear substantialistic structure. This points were however missing. In any textbook and in most books nothing was written from such a viewpoint, so that I had to find out them by examining original papers."

(2) The reader will see easily how "the most important points in making clear substantialistic structure" are grasped in the present book. We can cite some of them worthy of special mention, in order of chapter and section: the role played by W. A. Mickelson's atomistic theory of radiation in the derivation of the Wien distribution formula (Chap. 1, Sec. 2, §1, §2); the role which Einstein's introduction of quantum of light played in his establishment of the relativity theory (Chap. 1, Sec. 4); the importance of Lorentz's theory of electron in separating substantial molecules of matter from ether to make ether the medium of light only in vacuum, opening the way to the theory of material structure (Chap. 2, Sec. 3, §1); the first establishment of the real and substantial foundation of electron theory, by the determination of the specific charge of an electron by means of the Zeeman effect (Chap. 2, Sec. 3, §2); the important significance of the Nagaoka model of the atom in the process of the final determination of the model of the atom with nucleus, by Rutherford's study of various substantial models of the atom so far proposed (Chap. 2, Sec. 5, §2); the establishment of the model of the atom by Bohr's determination of the number of electrons in an atom of light elements, after Rutherford's determination of the model of the atom with nucleus and just before Bohr's application of quantum theory to atoms, by means of the loss of velocity of a charged particle passing through matter (Chap. 2, Sec. 6, §5).

As to the Wien distribution formula, Taketani points out that "the development of the theory of thermal radiation was accomplished by building theories
on suitable assumption from the atomistic viewpoint, before the results derived from the classical theory in strictly logical ways were shown to be incompatible with experiments (p. 14). As to the theory of light quantum, he points out that Einstein did “treat light as a group of molecules to which the kinetic theory is applied”, and “did not treated light with the idea of light as an oscillation of the ether”. Taketani then states that “we should say that the negation of the ether was not made for the first time in his paper on the theory of relativity, but was made for the first time in his paper on the theory of light quanta” (p. 51). Important is his indication that “the medium of ether was actively negated”, and “this did not arise from the passive viewpoint that, as is said in the Machist empiricism, the ether was abandoned because the ether was not observed” (p. 51). Taketani notes that Einstein avoided introducing directly the theory of light quanta just born into the theory of relativity according to the custom of German academism, but used the expression of “a given light complex” (p. 52).

As to Lorentz’s theory of electron, Taketani describes that by this theory electromagnetism “adopts the formulation of the interaction of the electromagnetic fields obeying the Maxwell equations with charged particles”, and “from the polarization and vibration of atoms of matter the optical and electromagnetic properties of matter were explained” (p. 80, p. 78). Lorentz’s theory, except the idea of ether in vacuum, plays an important role still in the present-day theory of matter. Taketani’s indication is important, that Lorentz’s way of thinking is different from “Kelvin’s idea of atoms being based on ether monism” (p. 75), but “is stepwise, and aims to give soundly substantiability to atoms and molecules, instead of trying to clarify at once the whole problem of atoms and molecules” (p. 78).

About the Nagaoka model of the atom, Taketani quotes Nagaoka’s words “I do not pretend that the system I propose represents a chemical atom, but by investigating the properties of an ideal atom, we may step by step aspire to a higher aim — the solution of natural atoms”, and highly appreciates them by stating that “it should be said that these words are indeed suggestive” (p. 181). As to the determination by Bohr of the number of electrons in an atom, Taketani attaches importance to its epistemological significance, and says as follows (p. 222) — “The hydrogen atom was shown to have only one electron. This finding brought the decisive difficulty of the stability and permanence of atoms to the classical theoretic treatment of the model of the
atom with nucleus. Bohr's theory of atoms based fundamentally on the determination of the model of such nature."

Now, let us note here a paper appeared afterwards which supports Take­
tani's analysis of the relation between the theory of light quanta and the theory of relativity. It is Max Born's lecture with the title "Physics and Relativity", given on Jul. 16, 1955 at the International Relativity Conference held in Berne (included in M. Born, "Physics in My Generation", Pergamon, 1956). In this lecture, Born cites Einstein's following words in Einstein's answer written in German in 1955 to Carl Seelig's question. "What was new of it, was the cognition that the significance of the Lorentz transformation went beyond the connection with the Maxwell equations and concerned the essence of space and time. Also the view was new, that the "Lorentz invariance" is to be a general condition for every physical theory. This was for me of particular importance, because I had really formerly recognized that the Maxwell theory does not represent the microstructure of radiation and for that reason is not generally maintainable." Born states then — "The last sentence of this is of particular importance. For it shows that Einstein's papers of 1905 on relativity and on the light quantum were not disconnected."

In Taketani's paper "On the Formation of Newton Mechanics", it is given as the second of the three forms of the movement from substantialistic to essentialistic stage, "the case in which a substance is dissolved into something functional, or conversely, a function has been grasped as a substance, good examples of this case being phlogiston and ether". But, as the analyses quoted above of Einstein's theory of light quanta and Lorentz's theory of electrons show, at the same time when ether was dissolved into a function, two substances of light quantum and material molecule appeared. The former is "an entirely new substance which requires a new logic", and corresponds to the third form of the movement from substantialistic to essentialistic stage in Take­
tani's theory. The latter corresponds to the first form of the movement in which "the introduction of substance leads at once to essentialistic theory", bringing forth Lorentz's theory of refractive index and dielectric constant. This does not mean, however, the field concerned of the theory of matter is completed with it, but means that from the interaction between electromagnetic fields and material molecules phenomena can be deduced. According to the specialty of the structure of material molecule and the interaction mentioned, different substances play their respective essential roles in different stages or phases of the deduction. In the theory of matter a substance does not necessarily mean a
The substance that governs the specific heat of a solid at low temperatures is the phonon, which is the quantum of elastic waves propagating through the solid. When Engels said that “life is the mode of existence of protein” (F. Engels, “Dialectic of Nature”), he knew that any protein consists of molecules, though he did not yet know that it consists ultimately of elementary particles. The later development of biology has made the content of Engels’s proposition more and more abundant.

Scientific cognition “develops through repeated rings” of phenomenological, substantialistic and essential stages (“Collected Works, 1”, p. 93), and therefore the three stages must not be taken in a fixed and schematic way. In this connection, it is also necessary to study enough S. Sakata’s “Philosophy of Stratum Character of Nature” (“Iwanami Series of Lectures on Philosophy 6”, Supplement No. 50, both being cited above, and “New Wind to Science”, Sin-Nippon Pub. Co., 1966). In Sakata’s paper “Elementary-Particle Theory and Philosophy” ([Kagaku, Apr. Issue, 1965]), he states that “the characteristic of physics in the 20th century consists first of all in the cognition of the stratum character of nature”, and that atomic physics in the 20th century has rediscovered dialectic view of nature, or, “the view that in the natural world there is an infinite series of strata of different qualities, from nebulae and the solar system, the larger ones, to molecules, atoms and elementary particles, the smaller ones, each of these strata being governed by laws proper to it, and the viewpoint that these strata are all the time in creation and disappearance, and in mutual relation and dependence as well, to constitute an interconnected world of nature”. Sakata states further that “Taketani’s theory of three stages has its objective ground on the dialectic of nature, namely on the multi-stratum structure of nature”. From the considerations described thus far, it is clear of itself that the so-called critique is quite meaningless, which has been sometimes raised against Taketani’s theory of three stages to claim that substance is unnecessary and matter is enough.

(3) In Preface to the present book, Taketani said “This book consists of Vol. I, Formation of Atomic Models, Vol. II, Formation of Quantum Mechanics and Vol. III, Logic of Quantum Mechanics”. However, Vols. II and III have not yet been accomplished. This is because he “has had no time to return to history”, being engaged on the one hand in contributing to the analysis and hence development of elementary-particle theory after the World War II, to the development of space physics, etc., as is described in Supplement No. 50 cited above and in his book “Present-day Theoretical Problems” (Iwanami Pub. Co.,
1968), and on the other hand in carrying out the practice to fulfill responsibility of scientist for society in the problems of resisting nuclear armaments and tests, of demanding safety of nuclear power plants and chemical industries, of defending people from pollution, etc., as is described in his “Collected Words 2, 3” and “Present-day Theoretical Problems”.

About the logic of quantum mechanics, Taketani analyzed its dialectic nature already in 1936 in his paper “Dialectic of Nature (On Quantum Mechanics)” (included in “Collected Works 1” and Supplement No. 50). In this paper, it was made clear how the logic by means of which the wave and particle natures of an electron or quantum of light are mediated from the essentialistic law of Schrödinger equation. This makes also the observation problem in quantum mechanics, about which the reader is referred to Taketani’s paper “Observation Problem in Quantum Mechanics” (included in “Present-day Theoretical Problems” and Supplement No. 50).

Taketani’s sharp analyses of the safety problem of nuclear power plants, and of the safety problem in general, are based on his theory of technology, which has been richly developed through his practice. Taketani’s theory of technology was born together with his theory of three stages, and is based on his dialectic grasping of the connection between the cognition and practices by human beings — the reader will remember that in his paper “On the Formation of Newton Mechanics” it was made clear that “science progresses by being governed by the three of technology, constitution of nature itself and mode of thinking”. As to Taketani’s theory of technology, the reader is referred to Taketani’s paper “Theory of Technology” (included in “Collected works 1”, p. 125) or the book “General Theory of Natural Science” by Taketani as an author and the editor (Keiso Pub. Co.). In the paper “Theory of Technology”, the reader will read Taketani’s famous proposition put forth in 1939 that “technology is the conscious application in practice by human beings of objective lawfulness”.

In the preface to the present book, Taketani said as follows — “As a work on the history of science, the present one makes only the fundamental study as an initial step. It is still necessary to make clear the relation of science to society and the interaction of science with general philosophical thought. I plan for a study on ‘the experimental foundation of quantum mechanics’ as to the former, and a study on ‘the role played by idealism and materialism in the formation of atomic physics’ as to the latter.” This is a statement that
must be understood in relation to the mentioned three factors governing the progress of science.

About the latter Taketani spoke recently to the present writer as follows — “This book treats the formation of quantum mechanics, but there are many problems before the appearance of quantum theory. Above all, the subjects from ancient times of Energetik and Atomistik, these I would like to make clear of their logical side, by treating them as philosophical concepts on the one hand, and by relating them to substantialistic problems on the other. Moreover, it is a very interesting problem to study how the logical side is related to problems with a social background in the history of thought. Among materialists only Engels and Lenin made such studies, and the rest were still satisfied to cite those studies. I intended to develop such a study concretely.”

In Taketani’s paper “Modern Interchange between Natural Science and Social Science” (1949, included in “Collected Works 5”), we see his analyses along this line woven here and there into the work.

About the former, Taketani told also to the present writer — “This is the problem of studying how far a technology coming available at an age may make correct cognition of nature possible, by means of its application to experiments. A work I made after the World War II on a modern subject of this kind is “Study of Elementary-Particle Theory III” (Iwanami Pub. Co., 1951). It was also one of my aims to make clear with what a background a technology comes out. I intended to made such a study in all its aspects as Amano had made in some form on the history of thermodynamics.”

(4) While the republication of this book has not been accomplished, “objections” were raised by some so-called historians of science to the important points of this book. Although those “objections” are so unfounded as to be unworthy of being treated, I would like to criticize them on this occasion because of the harmful influence they exert on students and people who are interested in the history of science, and also because we can derive a lesson from the errors included in them.

T. Hirosige wrote an explanatory article entitled “Where was Theory of Relativity born?” in the Journal of Physical Society of Japan (in Japanese), Jun. 1972. In this article Hirosige says, about the relation between Einstein’s theories of light quantum and of relativity analyzed in the present Vol. I, that Taketani’s theory “is based on the premise that the very ether was the crux of the problem, so that Taketani’s theory cannot but fail when this premise fails as is described above”.
How does Hirosige show that "this premise fails"? Hirosige takes the fact that Einstein's first paper on the theory of relativity begins with a paragraph of pointing out the unsymmetry seen in Maxwell's electrodynamics for moving bodies. Einstein takes, as an example, the electrodynamic action between a magnet and a conductor, and states that, whereas the observable phenomenon depends only on the relative motion between the magnet and conductor, there arises an electric field in the neighbourhood of the magnet, if the magnet is in motion, producing an electric current in the conductor, but there arises no electric field in the neighbourhood of the magnet, if the conductor is in motion, and an electromotive force is produced to give rise to an electric current in the conductor. Hirosige says that this paragraph of Einstein shows that "the argument is not developed from the problem of the influence of the motion relative to the ether".

Let us listen, however, to Einstein's words a little more. After the first paragraph Einstein continues to say in the second one as follows — "Examples of this kind, as well as the unsuccessful attempts to detect any motion of the earth relative to the 'light medium', lead to the supposition that no property of phenomena corresponds to the concept of absolute rest, not only in mechanics but also in electrodynamics, but rather the same electromagnetic and optical laws are valid, also for all coordinate systems for which the equations of mechanics are valid, as is already proved for the quantities of the first order".

As any one who reads this paragraph can understand, the argument in Einstein's paper is developed from the negation of the concept of absolute rest. As is well known, the Newton mechanics satisfies the Galilei principle of relativity, so that the mechanical laws which are valid in an inertial system are also valid in another inertial system as the same laws. Nevertheless, Newton supposed the space absolutely at rest in the universe. Therefore, the absolute rest cannot be detected by means of phenomena in the domain of the Newton mechanics. Now, when the ether, that is the medium of light waves, was supposed to fill up the universe, it came to be considered that relative velocities to the absolute rest could be detected by means of optical phenomena. And, when it was discovered that any light is an electromagnetic wave, the ether became the medium of electromagnetic field.

The Maxwell theory of electromagnetism does not satisfy the Galilei principle of relativity, and it includes the velocity relative to the ether. Therefore the laws depend on the velocity of the coordinate system concerned. The electromagnetic induction between a magnet and a conductor is such an example.
On the other hand, the velocity of the earth relative to the ether, which should be detected by means of phenomena of light, remained being undiscovered at all. Thus, Einstein supposes that the absolute rest may have nothing to do with not only mechanics but also electrodynamics, and that the same laws of electromagnetic phenomena may be valid for all inertial systems, since this is so in fact in the domain of quantities of the first order. Einstein does bravely "raise this supposition to the premise", introducing the principle of relativity.

In spite of these facts, Hirosige says even that "the problem of the relative motion to the ether was mentioned only as one of the individual matters which support the theory thus developed". The very thing that Einstein did for the first time was the negation of the ether. The coordinate transformation which Einstein derived from his standpoint is called the Lorentz transformation, because it had been derived before Einstein by Lorentz from the standpoint of his electron theory. As is described in (2) above, it is the analysis given by Taketani that Einstein negated positively the ether on the basis of his theory of light quantum.

Einstein's work in his first paper on the relativity starts from the critical reconsideration of the meaning of the measurement of space and time on the basis of the principle of definite light velocity, to construct "a simple and uncontradicting electrodynamics". "The introduction of a 'light ether' will prove to be superfluous, insofar as according to the view to be developed neither an absolutely rest space provided with special properties will be introduced, nor a velocity vector will be assigned to a point of the empty space, in which electromagnetic processes take place." The great work is accomplished, and the new electrodynamics shows itself with no relation at all to the ether. These things are treated by Hirosige's hands in the manner that "although Einstein stresses that as a result of the theory of relativity the ether (as the standard of the absolute coordinates) becomes unnecessary, he never says that the theory of relativity becomes inevitable because the ether is unnecessary".

Hirosige says that "there has been prevailing such an interpretation that Einstein's theory of light quantum aimed at explaining the photoelectric effect". However, on p. 43 of the present volume, it is properly described that Einstein took the Stokes law and the photoelectric effect, after having shown that a monochromatic radiation behaves as a group of energy quanta when its density was low enough, in order to investigate whether light consisted of energy quanta or not. Hirosige neglects this description.
Examinations of Both Models and the Determination of Atomic Model

(5) In the Oct. 1971 issue of "Sizen" (Nature), Eri Yagi wrote an article in Japanese entitled "Rutherford and Hantaro Nagaoka". In this article Yagi takes up, saying to make an "analysis of material", a post card from Bragg to Rutherford dated Mar. 11, 1911. Yagi says that, because in this letter there were certain words to mean "It seems that Nagaoka once made an attempt to assume a large positive central nucleus in the atom", while Rutherford reported on Mar. 7 in the same year "a short article containing the idea of the model of the atom with nucleus", "Rutherford's model seems to be independent of Nagaoka's model". Yagi describes a similar thing to this in her explanatory notes to "Theories of Atomic Structure" (Tokai Univ. Press, 1969), one volume in a series of the translations into Japanese of classical papers on physics, in which Taketani's present Vol. I is cited as an "example" of the "review" of Nagaoka's model.

However, how can it be guaranteed that all the things Rutherford read or heard on one day were the things he did for the first time. It could not happen, just like all the articles in today's newspaper are about the things some one knows for the first time. It can never be possible, therefore, to derive simply such a conclusion as Yagi's from Bragg's post card. If justice is administered in a way similar to Yagi's, there will arise so many dark judgments. It needs a certain evidence to assert with objectivity that Rutherford did not know Nagaoka's paper, because Phil. Mag. was Rutherford's "desk-side journal" as Nagaoka said ["Kagaku Asahi" (Asahi Science), Jan. 1950].

Rutherford wrote a letter dated Mar. 20, 1911, to Nagaoka, in which Rutherford stated as follows — "You will notice the structure assumed in my atom is somewhat similar to that suggested by you in your paper some years ago. I have not yet looked up your paper; but I remember that you did write on that subject." (Italics by the present writer) Because of this 'remembrance', Rutherford could not note in his paper any word to mean that he 'arrived at this structure independently of Nagaoka'. The independence in scientific achievements is nothing but the same right about priority.

Rutherford himself describes in his paper, quite differently from Yagi's saying, "From the point of view considered in this paper, the chance of large deflection would practically be unaltered, whether the atom is considered to be a disk or a sphere." (Cf. p. 221 of the present volume.) That is, Rutherford recognizes that the model for his analysis of scatterings of large deflection is essentially identical with Nagaoka's model. Nevertheless, Yagi says that Rutherford's "mention is very superficial" and "gives no consideration at all
about the mechanical stability of Nagaoka’s model”. However, how does Yagi see the fact, that Rutherford himself mentions that “the question of the stability of the atom needs not be considered, for this will obviously depend upon the minute structure of the atom, and on the motion of the constituent parts”?
(Cf. p. 209 of the present volume.)

Yagi states then, saying to make an “analysis from academic side”, that the central charge in Nagaoka’s model “is not demanded at all to be very small (smaller than $10^{-12}$ cm) as is in Rutherford’s model”, and that “the outer radius of the ring of Saturn, after which Nagaoka modeled his theory, is only about two times the radius of the equator of Saturn”. Nagaoka took his model after Saturn, and wrote also “In his lectures on electrons, Sir Oliver Lodge calls attention to a Saturnian system.” However, did Nagaoka write that his model was a miniature in a proportionally reduced scale of Saturn? Nagaoka defined clearly “the system” which he was “going to discuss” as is seen on p. 169 of the present volume, and stated then “The system differs from the Saturnian system considered by Maxwell in having repelling particles instead of attracting satellites.” This is to indicate that Nagaoka’s model has not a similar nature to the Solar system, but to the Saturnian system. The similarity between dynamical systems never means the proportionality in a reduced scale, but the dominance of the essentially same equation over the systems.

Nagaoka said, in his paper in “Kagaku Asahi” cited above, “I confined myself to give about the nucleus the view that it should be positively charged and its mass should be large compared with that of an electron.” This was just so. It did not need to assume a value of the radius of the central charged body at the time of having no clue of it, because the electric field due to the central charged body at a point outside the body is the same as that due to a point charge at the central origin. Important is Nagaoka’s view itself, which Rutherford appreciated as we have seen above. Rutherford’s study confirmed (not presented) for the first time the existence of a central nucleus in an atom, and made clear at the same time that its size is of the order of magnitude of at most $10^{-12}$ cm. On account of this, it is stated on p. 221 of the present volume that “Rutherford’s theory brings thus the complete determination to the question of the models of the atom” without or with nucleus.

Examinations of Both Models and the Determination of Atomic Model

one volume in the series cited above, and says in them even that "Rutherford's model is never opposed to Thomson's model, but can be seen as a variation of the latter". According to Hirosige, "the word 'the structure of the atom' in the title of Rutherford's paper should not hastily be understood to mean the problem of the atom being with or without nucleus".

Hirosige does not make any mention about the fact that Rutherford wrote clearly those words about Nagaoka's model we have seen above in (5). Even on account of this point, Hirosige's explanatory notes are improper. Even if we leave aside this point, Hirosige's remark is haphazard. Saying as in the quotation just above, he insists that Rutherford's model "is not of electrons rotating around the nucleus, but of a uniform distribution of opposite electric charges around the center of force which consists of electric charges concentrating in a small domain". Hirosige continues to say proudly that the central charge is "supposed for the moment to be plus or minus".

Rutherford considers in his paper, with the purpose of "examining theoretically the single encounters with an atom of simple structure, which is able to produce large deflections of α-particles, and then comparing the deductions from the theory with the experimental data available", the electrostatic potential for a point inside an atom containing a positive charge $Ne$ surrounded by a distribution of negative electricity $Ne$, "in order to form some idea of forces required to deflect an α-particle through a large angle". He calculated by means of the conservation law of energy the nearest distance from the center an α-particle could arrive at, finding values of the order of $10^{-12}$ cm, values much smaller than the radius of the atom. Therefore, "the field due to the uniform distribution of negative electricity may be neglected" (Cf. p. 209 of the present volume). If the negative electricity is "distributed in the form of corpuscles" as in the case of Nagaoka's model, "its effect is in general small compared with that due to the central field", since it is shown that "the chance of single deflections through a large angle due to this cause, if not zero, must be exceedingly small compared with that due to the central charge", because the mass, momentum and kinetic energy of the α particle are very large compared with the corresponding values for an electron, so that the α-particle cannot be deflected through a large angle by a close approach to an electron (Cf. p. 220 of the present volume). Because of this, Rutherford makes the mention about Nagaoka's model as is quoted in (5) above. The above considerations are also "in support of the assumption" of having neglected the effect of the negative electricity supposed to distribute throughout the atom.
In this way, on the assumption that large deflections of an \( \alpha \)-particle are caused by a single collision with the central charge, Rutherford calculates the probability of getting a given angle of deflection, and shows that the theoretical result agrees well with the experimental data given by Geiger and Marsden. Because the calculated probability is proportional to \((Ne)^2\), the same conclusion will be deduced even if the central charge is negative, i.e. \((-Ne)\), as long as this probability is measured. Hence, from the experiment of large deflection, the existence of the central charge can be confirmed, but it cannot be concluded whether it is positive or negative. However, Rutherford remarks that the reduction in the velocity of \( \beta \)-particles should be more marked with a positive center, and the high velocity of expulsion of \( \alpha \)-particles may be accounted for by the positive central charge.

Hirosige says furthermore that "a careful reading of Rutherford's paper will make it clear that even the possibility is reserved of the existence of charged particles scattered in the distribution of the charges of the opposite sign". There is a Japanese proverb "Seeing trees but not the forest". This saying of Hirosige cannot be remarked but as "Seeing leaves but not the tree". A glance of Rutherford's paper makes it clear that Rutherford considers the case of negatively charged corpuscles scattered in the atom as is described above, and says about the positive charge that the experimental evidence is not precise enough to negate the existence of satellites carrying a fraction of the positive charge (Cf. p. 220 of the present volume). In Thomson's theory of the scattering of \( \beta \)-particles (1910), the case of the positive charge divided into parts is discussed, too. But, because these parts distribute uniformly throughout the atom, this case does not alter the situation that there is no positive charge concentrated at the center. In short, in the series of Thomson's model of the atom by means of corpuscles (1897) discussed in Chap. 2 of the present book, Kelvin's model of the atom (1902), and Thomson's model of the atom (1903 ~), there is no positively charged nucleus supposed. The reader will not need more explanation to see how Hirosige's saying is nonsense that the title of Rutherford's paper "should not hastily be understood to mean the problem of the atom being with or without nucleus".

Hirosige says that "the structure of the atom" in those days means first of all "the problem of the number of inneratomic electrons", and Rutherford himself "investigates with the greatest care possible the number of inneratomic electrons". However, for Rutherford "the value of the central charge \( Ne \) is an important constant, and it is desirable to determine its value for
Examinations of Both Models and the Determination of Atomic Model

different atoms”. Rutherford discusses thus, in addition to the three kinds of experiments on the scattering of α-particles, the experiments of Crowther on the scattering of β-rays” (Cf. p. 215 of the present volume). Crowther’s experiments were considered to verify Thomson’s theory of multiple scattering of β-rays. However, Rutherford shows that “the main relations on the theory of compound scattering … holds equally well on the theory of single scattering”, and calculates the values of $N$ from Crowther’s experimental data. On the contrary, Crowther concluded from the same experimental data with the use of Thomson’s equation that the positive electricity distributed continuously throughout an atom, and the number of electrons in an atom was about three times its atomic weight (Cf. p. 207 of the present volume).

Rutherford uses the most pages for the discussion of Crowther’s experiment among the four kinds of experiments treated, because Crowther deduced the conclusion quite opposite to Rutherford’s conclusion. How can one see this as showing that “Rutherford pays his attention most to the study of the number of inneratomic electrons” as Hirosige says? No more of deception like that encountered at some street stalls. About the problem of the number of inneratomic electrons, the substantial structure of hydrogen atom is established by Bohr, and the way of applying the quantum theory to it is then sought, as we have seen in (2) above.

It is common to these so-called “historians of science” who find fault with the analyses given in the present volume, that they do not understand the very difficulty of making a really original step forward.

[The rest of the present explanatory notes is omitted, in which some similar examples in the field of the theory of elementary particles are treated.]
THE FORMATION AND LOGIC OF QUANTUM MECHANICS

Volume II

The Way to Quantum Mechanics

by

MITUO TAKETANI

MASAYUKI NAGASAKI

Translation from the Japanese Edition by
Masayuki Nagasaki

World Scientific
THE

FORMATION AND LOGIC

OF

QUANTUM MECHANICS

Vol. II

The Way to Quantum Mechanics
This page is intentionally left blank
THE FORMATION AND LOGIC OF QUANTUM MECHANICS

Vol. II
The Way to Quantum Mechanics

by
MITUO TAKETANI
MASAYUKI NAGASAKI

Translation from the Japanese Edition by Masayuki Nagasaki

World Scientific
New Jersey • London • Singapore • Hong Kong
Preface to Volume II

It was in 1948 when "The Formation and Logic of Quantum Mechanics; Vol. I — The Formation of Atomic Models" was published. This Vol. I was republished in 1972 by means of photographic printing from Keiso Pub. Co. Ltd. with explanatory notes by Nagasaki added to. Of the circumstances that delayed the publication of Vols. II and III, announced beforehand in the preface to Vol. I, descriptions were given in the above-mentioned explanatory notes.

Volume II, of which early publication was hoped, has now been completed under the title "The Formation and Logic of Quantum Mechanics; Vol. II — The Way to Quantum Mechanics" by the collaboration of Taketani and Nagasaki. The authors apologize to the reader for the delay of more than forty years after the publication of Vol. I.

The principal purpose of Vol. II is, as that of Vol. I, to make clear the logic of the development of scientific cognition. In Vol. I we treated, in Chap. 1 on the quantum theory of radiation, the contradiction between the theory of electromagnetic wave and the atomistic point of view, Planck’s theory of thermal radiation, and Einstein’s proposal of the theory of quantum of light, that is, the appearance of the contradiction between undulatory and corpuscular theories. We treated, in Chap. 2 on the formation of atomic models, the model of the atom without nucleus and the model of the atom with nucleus, and discussed various attempts to study atomic spectra according to the theory of electron and on the basis of the classical theory of electromagnetic wave.

In Vol. I descriptions were given of the process up to the confirmation by Bohr, after the experimental verification by Rutherford of the Nagaoka model of the atom with nucleus, of the fact that there is one electron in an
atom of hydrogen, the confirmation by which the model of classical oscillation for atomic spectra was brought to a deadlock. This means that, seen from Taketani's three-stage theory of the cognition of nature, the basis of the substantialistic stage in the theory to cognize atoms was obtained.

As to the three-stage theory a brief description was given in the concluding remarks of Vol. I, and a further description was added by Nagasaki in the explanatory notes in the republished edition of Vol. I. The three-stage theory of the cognition of nature was developed by Taketani, on the basis of his critique that only with Machist epistemology, according to which one obtains mathematical laws by systematizing given sensuous experience, one is powerless in practice.

In the first stage of the cognition of nature (the phenomenological stage), description of a phenomenon or that of experiential results is made. It is the stage of getting only the knowledge (or law) of the phenomenon without explaining the phenomenon by relating it further to other facts. In the second stage (the substantialistic stage), the substantialistic structure to bring forth the phenomenon is obtained, by the knowledge of which the description of phenomenon is systematized to be provided with lawfulness. This lawfulness does not mean, however, that after an event must necessarily follow another.

In the third stage (the essentialistic stage), the cognition deepens to the essence by the mediation of the substantialistic stage, so that the law of phenomenon is explained through the mediation of necessary motion of the substances under their mutual interaction. These three stages indicate logical stages, and should not be taken in a mechanical way because the respective logic appears in various forms in accordance with the reality.

In the present Vol. II we shall consider how atomic phenomena were elucidated by the application of quantum theory. The process will be treated that a substantialistic stage was advanced to be established, or, how contradictions arising one after another were pursued, systematized and made clear to proceed to the following essentialistic stage, a drama, so to speak, full of complications.

Regarding the theory of radiation, it was made clear that the Planck formula contained the fundamentally contradicting concepts of wave nature and particle nature, and this sharpened the logical inconsistency. Einstein introduced a probabilistic concept to settle the problem, but it made the difficulty more serious.

In the theory of atoms, Nicholson gave a clue to bring the Planck constant \( h \) into the theory based on the Nagaoka model of the atom with nucleus, setting
up the footing for the establishment of Bohr's famous quantum-theoretical model of the atom.

The Bohr model of the atom introduced $\hbar$ into the motion of inner-atomic electrons, by supposing some basic phenomenological assumptions. The start of efforts to bring mechanics into them was made by the presentation of Sommerfeld's quantum condition. In this way, the research of a new mechanics was advanced, and the spectra of light and X-rays emitted from atoms were elucidated more and more, to make clear the configuration of electrons in various complex atoms.

The Mendelev periodic table, which had been presented in the 19th century and played a very important role in chemistry, was entirely in the phenomenological stage as seen from the viewpoint of the theory of atomic structure, but proceeded to the substantialistic stage when it was then elucidated by the analysis of the configuration of inner-atomic electrons.

However, the attempt of making clear the problem of inner-atomic electrons by means of the quantum condition met with difficulties one after another, and in the process of making effort to settle them the characteristic of contradictions was made clear. On the other hand, the interconnection of the theory of quantum of light and the theory of electromagnetic wave was elucidated by the mediation of the quantum condition.

The final step to quantum mechanics was thus prepared. The content of Vol. II concerns the analysis of the process coming to this final step. In other words, it is to make clear the logic of the most interesting and complicated transition in which the way to the essential theory was developed by making substantialistic systematization.

As to this point, one will remember that the way from the introduction of ether in undulatory optics or in the theory of electromagnetic wave to the theory of quantum of light, was just in a similar stage. At the beginning of the 19th century the phenomenon of interference of light had been discovered by Young and Fresnel, and thereby the undulatory theory of light had been established. The wave concerned was supposed to be a wave in a medium called ether. In other words, a substance named ether was introduced. From the fact that light propagates through matter, as well as vacuum, it was concluded that light was an elastic wave in the ether pervading over vacuum and matter. Various studies were made of it on the basis of the classical theory of elasticity, but came to be confronted with unsolvable contradictions.
By the establishment of the Maxwell theory of electromagnetic wave, the theory of ether was brought to the next stage of making the nature of electromagnetism clear. When it became clear that a part of the role of ether in matter was due to the electric polarization of molecules in matter, there remained the problem of the propagation of light in vacuum. Various measurements were made in vain of the velocity of an observer moving in the ether in vacuum relative to it, and serious difficulties were thereby brought on. Ether was at last dissolved by Einstein’s theories of light-quantum and of relativity in 1905. However, this does not mean that all the substantial elements of light were dissolved. The substance of ether was transferred into the substance of electrically polarizable material molecules and the new substance of light quantum. In this way, the ether theory of light became temporarily an essential theory, which made the theory of light quantum containing new sharp contradictions start.

In the period treated in Vols. I and II, many attempts were made of getting a substantialistic image of atomic structure. However, it became clear that the world of atoms was accompanied by jumps in change, as far as this world was governed by the Planck constant $h$. Since classical images were based on the idea of continuity, there arose a large gap between them and the world governed by $h$. Therefore, those attempts were followed by the difficulty of how to find a route to climb a mountain surrounded by cliffs, so to speak. However, there arose some experimental results which could be explained well without the use of $h$. By means of them the climbing route connecting classical substantialistic images and the world of atoms was found.

The first of them is the success of Lorentz’s classical theory of electromagnetism for the normal Zeeman effect we treated in Vol. I. Because the result deduced with the use of the classical theory of electromagnetism as it stood remained to hold good in the world of atoms, it served to get the image of atoms. The second is the Rutherford scattering we treated, too. Also in this case, the result of the calculation by means of classical theory of the cross section as it stood was found to hold good in quantum theory. Because of this fact the model of the atom with nucleus was verified. The third is the calculation by means of the classical theory of the stopping power of matter for charged particles. As we described in Vol. I, by applying its result Bohr confirmed that the number of electrons in a molecule of hydrogen is two.

The fourth is Frank–Hertz’s experiment which we shall treat in the present Vol. II. The classical theory suffices for this experiment by itself, but it revealed
the existence of $h$ in the world of atoms as it confirmed the existence of discrete stationary states in the Bohr model of the atom. The fifth is Bohr’s correspondence principle. This was a generalization of the connection which could be taken to be valid between the classical and quantum theories as to some of the atomic phenomena in the case of large quantum numbers.

In the present book, we concentrate our aim on the investigation of the logic of the development of scientific study itself, and leave the description of social backgrounds, atmospheres of the times concerned, biographical elements, the so-called entertaining accounts, etc. to other authors. Letters and autobiographic recollections of famous scientists have been published in recent years. However, in the present book we carry forward, irrespectively of them, our analyses on the basis of papers published in academic journals. This is because it is necessary to prevent such descriptions as often seen in works on the history of science, that fall into stories laying stress on great scholars and tend to hero worship, so to speak, obscuring the course of the development of logic. The course mentioned can be made clear enough by following discussions exchanged in academic journals that are open to the public. On the other hand, the letters of the authors of those papers which gave some clues or hints to works of great scholars are not published, unless those authors are also great scholars.

At the beginning of each chapter of the present volume, we put a figure to give an easier understanding of our analysis of the way, how the complex development of logic was made under the mutual influence of researches. Each figure is not a flow chart of indicating an operation B to be done after an operation A, nor a diagram to show the way to a given destination. Researchers ought not to know where they will arrive at. They tackle problems with which they are individually confronted, making efforts to overcome contradictions. It sometimes gives a researcher clues for a new development to find contradiction in a work of another researcher, or to find significance in its result to which the other researcher himself does not attach importance. The present book is written from the viewpoint of those researchers who advanced their studies.

The method of analysis used for the present book of following published papers only is made possible partly because of the situation of the academic world in those days. Different from the present-day academic world, the number of researching circles as well as researchers was small, and papers of other researchers were comparatively fairly cited. There was the consciousness of fellows among researchers. This situation was utterly different from the
present-day one. Nowadays, papers come to have the character of merchandise, on account of the great expenses of making study. Papers are not fairly cited. The times before the World War II were the ones in which European and American researchers read carefully and cited those papers which were written by H. Yukawa, S. Sakata, Taketani, and others, who had not been in Europe and U.S.A. and had no acquaintance with European and American researchers, and which were published in a journal of Japan, internationally a local region in those days. As a typical example to show that the situation has become different after the War, we may mention the case of the Sakata model of the compositeness of elementary particles proposed in 1955, which has not been duly cited by foreign researchers.

To get an aid for understanding the present volume, general readers may consult S. Tomonaga "Quantum Mechanics, Vol. I" (Ginza Pub. Co., 1948). Its first publication was unexpected nearly at the same time after the War as that of Vol. I of the present book. The book by Tomonaga seems at a glance to describe things in the historical order, but it is not a book on the history of science. However, it is written in a clear-cut way as a textbook of quantum mechanics, because it explains and arranges materials of the old quantum mechanics, such as the Boltzmann principle, Ehrenfest's adiabatic hypothesis, Sommerfeld's quantum condition and Bohr's correspondence principle, on the basis of a fundamental way of thinking established later, as to these materials except the Boltzmann principle, as a result of the investigations of quantum theory.

In the present book, we shall make clear, oppositely to the viewpoint of textbooks of quantum theory, that these principles were obtained as a result of many investigations after many meanders. The present book tries to describe, from the respective viewpoints of many physicists of those times, the drama in the history of their closing in upon a gigantic object from various sides, and to elucidate the logic of their advances.

For the purpose of the present Vol. II, it is important to follow the consciousness of problem and development of logic in each original paper in accordance with the expressions in the original paper. However, because the translation itself of original papers is not the purpose of the present book, and because, if quotations from the original papers are often inserted into their description, the proceeding of description is interrupted and the reading becomes troublesome, we have made efforts to weave, with the use of quotation marks, decisive phrases or clauses of the original papers into our descriptive sentences.
The book that we have most reliably consulted in writing the present Vol. II is the 1919, 1922 and 1924 editions of "Atombau und Spektrallinien" by S. Sommerfeld. He was a physicist who struggled to elucidate atomic phenomena and played a central role in the period treated in our Vol. II. Though the above-mentioned book by Sommerfeld was not written as a book on the history of science but as a voluminous textbook, it got together the results of investigations up to the time of each edition. It is very helpful to follow these editions in order, to understand the change in the cognition and consciousness of the problems in that period. Besides this book, "Ursprung und Entwicklung der älteren Quantentheorie" by A. Rubinowicz [Handbuch der Physik (1933, Springer), Band 24, Kap. 1] gives references to important papers and summarizes the matter concerned compactly affording a general view. Rubinowicz was a physicist who contributed to the research under Sommerfeld in that period, which will be treated in the present volume.

There are various types of books on the history of science. Some describe the history of science with respect to society, epoch or thoughts as its background. In this type, those books which describe the history of science in relation to productive technology are included. And there are such books as describe works of scientists arranged in the order of achievement and their careers or the circumstances among them. These books conveniently served as handbooks, so to speak.

As to the history of quantum mechanics, we have in Japan the pioneering and outstanding book "The Origin of Theory of Thermal Radiation and Quantum Theory" by Kiyoshi Amano, and, further, the book "The History of Quantum Mechanics" by the same author. In the former its relation to the foundation of the productive technology is well written, and in the latter the atmosphere of the academic world in those days is well described. These make both books excellent works. Furthermore, the explanation in them are to the point so as to be understandable by beginners.

M. Jammer's "The Conceptional Development of Quantum Mechanics" (1966, McGraw-Hill), of which there is a translated edition in Japanese, is comparatively comprehensive and is convenient as a handbook. Jammer makes use a number of times of the sayings, in later years, of researchers heard by him. However, this makes his descriptions apt to look back at the process of development from the viewpoint arrived at as a result of the development of quantum mechanics. This Jammer's book gives the name and volume number
of the journals in which many papers were published, together with their titles, and serves conveniently as a bibliography of literature.

Parts 1 and 2 of Vol. I of J. Mehra and H. Rechenberg’s "The Historical Development of Quantum Theory" (1982, Springer-Verlag) treat the period of 1900–1925, nearly the same as the one taken in our present book, and gives the careers of many researchers besides a list of many papers. They give to their Vol. I the main title “The Quantum Theories of Planck, Einstein, Bohr and Sommerfeld”. In spite of the great roles played by these four physicists in the development of quantum mechanics, there is a risk of getting into hero-worship by summarizing this period with such a title. They make much use of the letters of great scholars. This is not however an appropriate way of following the course of the development of logic, as we have mentioned above. Some characteristic mistakes of Jammer and of Mehra and Rechenberg will be pointed out in our text.

In writing the present volume, we were helped by Prof. M. Konuma when he was at the Institute of Fundamental Physics, Kyoto University, to get some old papers. We have utilized the inter-university cooperation system for exchange of literature. In utilizing this system, we have asked the assistance of the members of the reference room of the Library, Rikkyo University, and of Miss T. Ohsuga of the Laboratory of Theoretical Physics, Department of Physics, Rikkyo University, to whom we are very thankful.

To the members of the Keiso Pub. Co., in particular to Miss M. Ishii, we are indebted for their patience in waiting for a certain period from the time of planning of the present volume on account of Taketani’s and Nagasaki’s circumstances.

Mituo TAKETANI
Masayuki NAGASAKI
June 16, 1990
Contents

Preface to Volume II v

Volume II. The Way to Quantum Mechanics 1

Chapter 1. Difficulties in Radiation Theory 3
1. Difficulties seen from Statistical Heat Theory 6
   §1. Phenomenological character of Planck’s resonator 6
   §2. The contradiction with the equipartition law of energy 12
   §3. Difficulties about Planck’s complexion 15
2. Molecular Theoretical Significance of the Planck Theory 19
   §1. Atomic vibration model for specific heat of solid body 19
   §2. Theory of specific heat of solids in terms of the quantum theory of sound 21
3. Conflict between the Wave and Particle Natures 25
   §1. Fluctuation of radiation energy 25
   §2. Fluctuation of radiation pressure 30
   §3. Conflict between the concepts of light wave and of light quantum 32

Chapter 2. The Quantum of Action and Atomic Models 37
1. Theoretical Significance of the Quantum of Action 40
   §1. The quantum of action and fundamental physical constants 40
§2. The Thomson model of the atom and the quantum of action .................................. 43
§3. The quantum of action and periodic and aperiodic molecular processes ................. 46

2. The Combination Rule of Spectral Lines ................................................................. 55
§1. Ritz's theory of spectral lines ................................................................. 55
§2. Ritz's "atomic magnetic field" model ...................................................... 60
§3. Ritz's combination rule .............................................................. 63

3. The Quantum of Action and the Angular Momentum ................................................ 67
§1. Nicholson's theory of spectra .............................................................. 67
§2. The quantum nature of angular momentum ............................................. 74

4. The Bohr Model of the Atom ................................................................. 80
§1. Characteristic X-rays and frequencies of atomic electrons ......................... 80
§2. The quantum of action and the fundamental assumptions in the Bohr model ...... 82
§3. The spectral formula and the angular momentum ..................................... 87
§4. The permanent state and the angular momentum ....................................... 93
§5. The structure of atoms and molecules ................................................ 95
§6. Moseley's verification of the Bohr model ............................................ 99

Chapter 3. The Quantum Condition, Transition Probability and Correspondence Principle 103
1. The Quantum Condition and Atomic Model .............................................. 107
§1. Wilson's quantum condition ............................................................ 107
§2. Ishiwara's quantum condition and the spectral formula ....................... 110
§3. The quantum theory of hydrogen atom due to Sommerfeld .................. 114

2. The Quantum Condition and the Multiplicity of Spectral Lines .................. 119
§1. The quantization of direction .............................................................. 119
§2. The relativistic fine structure of spectral line ..................................... 123
§3. The limitation of the quantum condition .............................................. 127
§4. The necessity of selection rules ............................................................ 128
§5. The quantum condition and the Stark effect ......................................... 132
§6. The Stark effect and Planck's quantum condition .................................. 135
§7. The quantum condition and the Zeeman effect ...................................... 140

3. The Transition Probability ................................................................. 144
§1. Two trends of thinking — energetics and atomism .................................. 144
§2. Understanding of the radiation formula in terms of transition probabilities ........................................ 147
§3. Light-quantum theoretical radiation and molecular recoils .................................................................. 151
§4. Demand for a quantum-natured theory of radiation ............................................................................ 155
4. The Correspondence Principle .............................................................................................................. 158
   §1. Angular momentum and the selection rule ......................................................................................... 158
   §2. The correspondence principle and selection rule ............................................................................... 160
   §3. The adiabatic hypothesis ...................................................................................................................... 164
   §4. The correspondence principle and adiabatic hypothesis .................................................................. 166
   §5. The correspondence principle and intensities of spectral lines ....................................................... 169

Chapter 4. Theory of Atomic Structure and Spin of Electron 175
1. The Anomalous Zeeman Effect and Half-Integer Quantum Numbers .................................................. 179
   §1. Inner quantum numbers and half-integer quantum numbers .............................................................. 179
      §1.1 Doublet and triplet terms and inner quantum numbers .............................................................. 179
      §1.2 Anomalous Zeeman effect of doublet and triplet terms and half-integer quantum numbers .......... 181
   §2. Half-integer quantum numbers and anomalous coefficients of magnetic moment .......................... 184
      §2.1 Half-integer quantum number and "atomic body" .................................................................. 184
      §2.2 The generalization of $g$-factor and the anomalous coefficient of magnetic moment .................. 186
   §3. The magnetomechanical anomaly ....................................................................................................... 189
   §4. Difficulties about the result of Stern-Gerlach's experiment ............................................................... 192
2. The Shell Structure in the Configuration of Atomic Electrons ............................................................... 196
   §1. The valence and the configuration of atomic electrons ................................................................. 196
   §2. Bohr's theory of atomic structure .................................................................................................... 199
   §3. X-ray spectra and structure of atoms ............................................................................................... 204
   §4. Stoner's revision of Bohr's configuration of atomic electrons ............................................................ 209
3. The Spin of Electron and the Exclusion Principle .................................................................................. 212
   §1. Pauli's analysis of the anomalous magnetic moment of the $K$-shell and the double-ness of the electron ...... 212
§2. The shell structure of atoms and the exclusion principle of electrons ........................................... 215
§3. The introduction of spin .................................................. 218

Chapter 5. The Interconnection of Wave- and Particle-Natures 227
1. Grasping of Wave Phenomena by Means of the Theory of Light Quantum ........................................... 231
   §1. The theory of light quantum and the Doppler effect ........ 231
   §2. The theory of light quantum and the Compton effect ........ 234
   §3. The theory of light quantum and the Bragg formula for X-ray reflection ........................................... 239
2. Light Quantum Theory and de Broglie’s Wave Theory ........ 244
   §1. Quantum condition and Brillouin’s theory of inner atomic elastic wave ........................................... 244
   §2. From the theory of light quantum to de Broglie’s wave theory ....................................................... 246
   §3. de Broglie’s wave and the diffraction and interference of particle .................................................... 249
   §4. Theory of light quanta and Bose-Einstein statistics .......... 252
3. Bohr-Kramers-Slater’s Theory and the Negation by Compton-Simon’s Experiment ................................. 262
   §1. Slater’s idea of virtual radiation ................................... 262
   §2. B-K-S theory of statistical conservation of energy-momentum ....................................................... 266
   §3. Compton-Simon’s experiment (Confirmation of the energy-momentum conservation in an elementary process) ... 268
4. Theory of Dispersion of Light .......................................... 272
   §1. Transition probabilities and the theory of dispersion of light ................................................ .......... 272
   §2. The theory of light quanta and the dispersion theory ........ 277

Postscript 285
Volume II

The Way to Quantum Mechanics
This page is intentionally left blank
Chapter 1

Difficulties in Radiation Theory
Michelson (1887)
Assuming the period of radiation $\tau \approx 1/u$ and the Maxwell distribution for the velocity of molecule $u$, got
$$\lambda_n^2 T = \text{const.}$$

Wien (1890)
Putting the wavelength $\lambda = g(u)$, and thus the exponent of the Maxwell distribution $-f(u)/T$, arrived at the Wien formula.

Planck (1900)
Assuming $dS/dU \approx 1/U^2$, derived from the equilibrium condition
$$U = (e^{\gamma/8\pi v^2}) \rho(v),$$
the Wien formula.

Planck (1900)
Putting $dS/dU \approx 1(U+\beta)$, arrived at the Planck formula.
Introduced the number of complexions and energy quantum.

Einstein (1905)
From the entropy of radiation obeying the Wien formula, proposed the hypothesis of light quantum.

Wien (1893)
The displacement law: for the wavelength $\lambda_m$ at which the radiation density is max.,
$$\lambda_m T = \text{const.}$$

Rayleigh (1905)
Assuming the law of equipartition of energy for long wavelengths, arrived at the Rayleigh-Jeans formula.

Jeans (1905)
Assuming thermal equilibrium for the case period of radiation > collision time, derived for long wavelengths the Rayleigh-Jeans formula.
Einstein (1906, 1907)
If the state density of resonators $w(U)$ is ordinary, $U = kT \rightarrow \text{R-J formula}$; if $0$ for $U \neq nhv$, $U = \text{Planck's value} \rightarrow \text{Planck formula}$. Applying Planck's $U$ to atoms of a solid, got the theory of specific heat.

Einstein (1909)
Pointed out the contradiction between Planck's complxions with the statistical theory, and the necessity of unification of wave and particle natures of radiation.

Lorentz (1910)
Gave paradoxes of light wave and light quantum.

Einstein (1911)
Derived $v$ of atomic oscillation form the compressibility of solid. Pointed out the spread of $v$ due to motion of adjacent atoms.

Debye (1910)
Applied Planck's formula to sound waves to get the specific heat of slids.

Jeans (1908)
Proposed the non-thermal equilibrium of cavity radiation of short wavelengths, by analogy with sound waves in a vessel.

Debye (1910)
Assuming the partition of energy of cavity radiation in units of $h\nu$, and using Planck's complxions derived Planck's formula.

Notes: 1) The numbers above a frame indicate Sec. and Subsec. in the present Chap, eg 1-1 indicates Vol. 1 - Chap. 1.
2) $U, S$ are the energy and entropy of a resonator, $p(v)$ the energy density of radiation, $v$ being the frequency.
Chapter 1

Difficulties in Radiation Theory

Theme in Chapter 1

The Planck theory of radiation contained various difficulties. As is described in Chap. 1 of Vol. I, Einstein pointed out that the Planck theory of radiation was contradictory to both the electromagnetic theory and the statistical heat theory, and that the hypothesis of the light quantum was brought into it.

The contradiction was contained in the phenomenological character of the resonator in the Planck theory. Though there were some attempts to defend the electromagnetic theory and the equipartition law of energy, the contradiction became intensified when it was clarified that there existed the energy quantum in the thermal oscillation of molecules in solid bodies.

It was further made clear that Planck’s distribution formula of radiation was to indicate the existence of the wave and particle natures of radiation, and then the necessity of a new radiation theory that would harmonize this duality was emphasized by Einstein, but the concepts of electromagnetic wave and light quantum were in so sharp a contradiction that there arose various paradoxes.

1. Difficulties seen from Statistical Heat Theory

§1. Phenomenological character of Planck’s resonator

As is already mentioned in Chap. 1, Vol. I of this book, the Planck theory of radiation aims to get the energy density of radiation from the average energy of the resonators which are in a thermal equilibrium with the radiation, without treating the radiation directly. The Planck theory is based on the relation

\[ U_\nu = \frac{c^3}{8\pi\nu^2} \rho_\nu, \tag{1} \]

which he derived for the energy \( \rho_\nu d\nu \) of the radiation in the frequency interval from \( \nu \) to \( \nu + d\nu \), and the average energy \( U_\nu \) of the resonators with the frequency \( \nu \) (\( c \) being the light velocity).
Einstein pointed out in his 1906 paper "On the Theory of the Generation and Absorption of Light,"\(^1\) as is mentioned also in Chap. 1, Vol. I, that the statistical heat theory gives \( U_\nu = kT \) at the temperature \( T \) (with the Boltzmann constant \( k \)), thus resulting in Rayleigh-Jeans' distribution formula for \( \rho_\nu \). He showed further, in his paper in 1907 "The Planck Theory of Radiation and the Theory of Specific Heat,"\(^2\) that if the energy of the resonator is assumed to take only such values that are integer multiples of \( h\nu \), instead of continuous values, \( U_\nu \) becomes

\[
U_\nu = \frac{\sum n\nu e^{-n\nu/kT}}{\sum e^{-n\nu/kT}} = \frac{h\nu}{e^{h\nu/kT} - 1},
\]

(2)

with which Eq. (1) gives us Planck's distribution formula for \( \rho_\nu \). If the energy of the resonator is supposed to change only with a jump, the usual electromagnetic theory could not be applied to the derivation of Eq. (1). On the other hand, if Eq. (2) is supposed to be valid, \( U_\nu \) becomes very smaller than \( h\nu \), and only a few resonators would have energies different from zero, in the domain of high \( \nu \) where Planck's distribution formula coincides with Wien's. Einstein pointed out in this way the contradictions involved in the Planck theory of radiation.

Einstein considered however that these contradictions do not negate the Planck theory of radiation, but indicate that the hypothesis of the light quantum is brought into the Planck theory of radiation. The hypothesis of the light quantum was the one proposed by Einstein to grasp the meaning of Wien's distribution formula.

By the way, the Planck theory of radiation began with his attempt to lay the foundation of Wien's distribution formula.\(^3\) Here for the sake of later discussions, let us look back upon how the contradictions pointed out by Einstein were brought into the Planck theory.

Michelson, in his 1887 paper "Theoretical Essay on the Distribution of Energy in the Spectra of Solids,"\(^4\) substituted, the thermal oscillation of each atom in a solid with the going-around motion of velocity \( u \) inside "the sphere of displacement" of radius \( a \), centered at the equilibrium point of the atom, and supposed that the radiation with the period \( \tau = 4a/u \) was most probably

---


\(^{2}\)A. Einstein, ibid. 22 (1907), 180.

\(^{3}\)Cf. Sec. 2, §3, §4, Chap. 1, Vol. I.

\(^{4}\)V. A. Michelson, Journ. d. Phys. 6 (1887), 467.
excited and had the highest intensity. He applied the Maxwell law of velocity distribution to \(u\), and assumed a factor of a power in \(1/\tau\) for the intensity of radiation. Making use of the law of the total intensity of radiation, called Stefan-Boltzmann’s law nowadays, he arrived at a distribution formula for the energy density \(\rho_\lambda\) of radiation as a function of the wave length \(\lambda\).\(^5\) \(\rho_\lambda\) so obtained decreases rapidly at short wave lengths, on account of the exponential factor in the Maxwell distribution.

The \(\rho_\lambda\) of Michelson becomes maximum at a wave length \(\lambda_m\) given by \(\lambda_m^2T = \text{const}\), because of the relation \(\tau \propto 1/u\) and the Maxwell distribution for \(u\). Since this does not agree with Wien’s displacement law,\(^6\) Wien assumed the form \(\rho_\lambda = F(\lambda) \exp\{-f(\lambda)/T\}\), in his paper in 1896 “On the Energy Distribution in the Emission Spectrum of a Black Body,”\(^7\) by supposing a more general relation between \(u\) and \(1/\tau\), i.e., between \(u\) and \(\lambda\), and obtained Wien’s distribution formula for \(\rho_\lambda\), by determining the functional forms of \(F(\lambda)\) and \(f(\lambda)\) with the use of Wien’s displacement law and Stefan-Boltzmann’s law. The factor of the exponential function in Wien’s distribution formula comes also from the Maxwell distribution. Wien considered that closer investigations of the atom as the source of radiation would be left with hypotheses which “will be useful for further improvement of the molecular theory”. That is, Wien tried to proceed stepwise on account of the lack of substantialistic knowledge of atoms, though he took the standpoint of the atomic theory.

Planck wrote, in his 1900 paper “On Irreversible Phenomena of Radiation,”\(^8\) “The progressive knowledge of the electromagnetic nature of the thermal radiation makes it an urgent task to grasp or possibly prove purely electromagnetically the second law of the heat theory in its application to the radiated heat”. Planck aimed at the foundation of Wien’s distribution formula, on the basis of laws already established in thermodynamics and electromagnetics. For this purpose, Planck considered the case of a resonator, which is the simplest electromagnetic system to emit and absorb radiation. According to him, “the emission of thermal radiation is regarded as caused by the sending out of electromagnetic waves from the side of certain elementary resonators, which may be seen to be in some connection with the ponderable atoms of the radiating body”.

\(^5\)Cf. the first equation on p. 13 of Vol. I.
\(^7\)W. Wien, Wied. Ann. 58 (1896), 662 [the same as Ref. 9) on p. 12 of Vol. I].
\(^8\)M. Planck, Ann. d. Phys. 1 (1900), 69.
Now, in his study in this paper, Planck puts the equation of motion for the electric dipole moment of the resonator $f$, oscillating with frequency $\nu$ in the direction of the $z$-axis under the radiation, as

$$\frac{d^2 f}{dt^2} + 2\sigma \nu \frac{df}{dt} + 4\pi^2 \nu^2 f = \frac{3c^3 \sigma}{4\pi^2 \nu} Z,$$

and expresses its energy $U$ as

$$U = \frac{1}{2} K f^2 + \frac{1}{2} L \left(\frac{df}{dt}\right)^2; \quad K = \frac{16\pi^4 \nu^3}{3c^3 \sigma}, \quad L = \frac{K}{4\pi^2 \nu^2}. \quad (4)$$

Here, $Z$ is the $z$-component of the electric field of the radiation, and $\sigma$ denotes the logarithmic damping rate of the resonator such that the energy of its free oscillations decreases by the factor $\exp(-2\sigma)$ per period owing to the emission of radiation ($\sigma$ being assumed to be not large). In this way, Planck does not give explicitly the mass and charge of the resonator, and stresses on its function of emitting and absorbing radiation.\(^9\)

Because of radiation damping, the frequency of the radiation emitted and absorbed by the resonator distributes around $\nu$ with a width of the order of $\nu \sigma$. By making use of the Fourier integral, Planck expresses $Z$ as

$$Z = \int_0^\infty d\nu'/C_{\nu'} \cos(2\pi \nu'/t - \theta_{\nu'}),$$

and then takes the average of $Z^2$ over a time interval that is short compared to the time passed but long compared to the period $\tau$ of the resonator. He calls this average value “the intensity of exciting oscillation” and denotes it as $J_\nu$. For the integrals with respect to $\nu'$ and $\nu''$ in the expression for $J_\nu$, Planck assumes that, putting $\mu = \nu' - \nu$, the “quantities rapidly varying” with $\nu'$, such as

$$C_{\mu+\nu'}C_{\nu'} \cos(\theta_{\mu+\nu'} - \theta_{\nu'}) \quad \text{and} \quad C_{\mu+\nu'}C_{\nu'} \sin(\theta_{\mu+\nu'} - \theta_{\nu'}),$$

\(^9\)Equation (3) used by Planck can be obtained as follows. Consider a resonator of mass $m$, electric charge $q$, and frequency $\nu$ under oscillation forced by the electric field $Z$. Take $f = qz$ instead of the displacement of $z$ of the resonator, and express the energy loss of the resonator owing to the emission of radiation as that due to friction. Eliminating $m$ and $q$ with the use of $2\nu = 8\pi q^2 / 3mc^2$, the ratio of the energy loss per unit time, he gets the final result.
may be replaced with some “slowly varying quantities” $A_\mu$ and $B_\mu$, respectively. He calls the radiation with this condition satisfied “the natural radiation”. $J_\nu$ changes slowly with time.

Inserting Eq. (5) into Eq. (3), Planck solves it under the assumption that, after a time long compared with the damping time $\tau/\sigma$, $f$ does not depend on the initial condition. Making use of this solution to calculate $U$, he applies the condition for “the natural radiation” to get the time average $U_\nu$ of $U$ similar to $J_\nu$ of $Z^2$, and finds the equation

$$\frac{dU_\nu}{dt} + 2\sigma \nu U_\nu = \frac{3c^3\sigma}{16\pi^2\nu^2} J_\nu.$$  \hspace{1cm} (6)

When the steady state is reached, $U_\nu$ becomes $U_\nu = (3c^3/32\pi^2\nu^2)J_\nu$, from which Eq. (1) is obtained through the relation $\rho_\nu = (3/4\pi)J_\nu$ between $\rho_\nu$ and $J_\nu$.

Planck then takes a completely phenomenological standpoint to “define” the entropy $S_\nu$ per resonator as

$$S_\nu = -\frac{U_\nu}{a\nu} \log \frac{U_\nu}{eb\nu},$$  \hspace{1cm} (7)

where $a$ and $b$ are constants and $e$ is the base of the natural logarithm. From this definition he calculates $U_\nu$ as a function of the temperature $T$ with the use of the thermodynamical relation $dS_\nu = dU_\nu/T$, and gets Wien’s distribution formula through Eq. (1).

Planck then tries, in his paper “The Entropy and Temperature of Radiated Heat,”\textsuperscript{10} to give reasons for “the definition” of $S_\nu$ given above. He considers that if the value of $U_\nu$ is changed by $\Delta U_\nu$ from its stationary value, the entropy of the system composed of the resonator and radiation will change by $dS_t$ in a short time interval $dt$, and the value of $U_\nu$ will change by $dU$, too, in the same interval. Making use of the calculations in the preceding paper quoted above, Planck shows that $dS_t$ is given by

$$dS_t = dU \Delta U \frac{3}{5} \frac{d^2U_\nu}{dS^2_\nu},$$  \hspace{1cm} (8)

and that $dU$ and $\Delta U$ are of opposite signs. Since $dS_t > 0$, he rewrites this as

$$dS_t = -dU \Delta U f(U_\nu),$$  \hspace{1cm} (9)

\textsuperscript{10}M. Planck, Ann. d. Phys. 1 (1900), 719 [the same as Ref. 14), Chap. 1 Vol. I].
putting
\[
\frac{3}{5} \frac{d^2 U}{dS^2} = -f(U). \tag{10}
\]

He next considers \(n\) resonators of the same character, and supposes that they behave independently and identically in \(dt\). Indicating the quantities under consideration related to the \(n\) resonators with the suffix \(n\), he puts \(U_n = nU, \Delta U_n = n\Delta U\) and \(dU_n = ndU\). Planck think that \(dS_t\) for the \(n\) resonators is given by \(-dU_n \Delta U_n f(U_n)\), which is obtained from Eq. (9) by replacing \(dU\) etc. respectively with \(dU_n\), etc. Planck further think that this is equal to \(n\) times Eq. (9), because he considers that the \(n\) resonators behave independently. That is, he puts
\[
dU_n \Delta U_n f(U_n) = ndU \Delta U f(U), \tag{11}\]
from which the relation \(nf(nU) = f(U),\) or \(f(U) \propto 1/U\), is deduced. Putting, therefore,
\[
\frac{d^2 U}{dS^2} = -\frac{\alpha}{U}, \tag{12}
\]
and integrating, Planck obtains an expression for \(S\) equivalent to Eq. (7). \(\alpha\) turns out to be equal to \(1/au\) from Wien’s displacement law.

In this way, Planck gave his reasons for Wien’s distribution formula. However, since it became necessary to modify it for the region of long wave lengths to meet the experimental results, he put in his paper “On an Improvement of Wien’s Spectral Formula,”\(^{11}\) instead of Eq. (12), the relation
\[
\frac{d^2 S}{dU^2} = -\frac{\gamma}{U(\beta + U)}, \tag{13}
\]
supposing that “it might be possible that the left-hand side expression [of Eq. (11)] does not possess in general the meaning ascribed before” by him. He then calculated \(S\) with a procedure similar to that just mentioned above, and arrived at Planck’s distribution formula.\(^{12}\)

As was shown by Einstein, in the region of short wave lengths Planck’s distribution formula coincides with Wien’s distribution formula, which should be grasped by means of the light quantum theory. From a comparison of Eq. (12)


\(^{12}\)Cf. Sec. 2, §7, Chap. 1, Vol. I.
and Eq. (13), one sees that $\alpha$ corresponds to $\gamma/\beta = k/\hbar \nu$, and thus Eq. (12) includes the Planck constant $\hbar$ substantially. When Planck put Eq. (11), which seemed self-evident to him, he made the very jump to the quantum theory without recognizing it. Planck, who aimed to found Wien's distribution formula on the theories of thermodynamics and electrodynamics, instead came to leave the classical theory, on account of the phenomenological nature mentioned above of the resonator he considered.

§2. **The contradiction with the equipartition law of energy**

Planck arrived at Wien's and Planck's distribution formulae, by assuming functional forms for the temperature dependence of the average energy of the resonator $U_{\nu}$, without treating directly the energy density $\rho_{\nu}$ of the radiation. From the fact that the cavity radiation is due to electromagnetic oscillations, it would be possible to treat $\rho_{\nu}$ directly, by applying the equipartition law of energy to these oscillations. A study from this standpoint was done, as is mentioned in Sec. 2, §5, Chap. 1, Vol. I, by Rayleigh in 1900 as a criticism against Wien's distribution formula. Jeans supplemented Rayleigh's thinking and tried to "defend" the equipartition law of energy.

About Wien's distribution formula Rayleigh considers, in his paper "Remarks upon the Law of Complete Radiation,"\(^{13}\) that although "viewed from the theoretical side the result appears to be little more than a conjecture" to him, "it is supported upon general thermodynamic grounds by Planck," and discusses that "speculation upon this subject is hampered by the difficulties which attend the Boltzmann–Maxwell doctrine of the partition of energy." He supposes that "although for some reason not yet explained the doctrine fails in general, it seems possible that it may apply to the graver modes," and puts

$$\rho_{\nu} d\nu = c_1 T \nu^2 e^{-\nu^2/T} d\nu,$$

making use of the analogy of the vibration of air in a cube to the electromagnetic oscillation of the cavity radiation.

In this equation, the factor $T$ comes from the fact that the average energy of each oscillation mode is proportional to $T$, and the factor $\nu^2 d\nu$ is due to the fact that the number of the oscillation modes in the frequency interval between $\nu$ and $\nu + d\nu$ is proportional to $\nu^2 d\nu$. The exponential factor is to exclude high oscillation modes, and has no theoretical ground. In this relation,

\(^{13}\)Lord Rayleigh, Phil. Mag. 49 (1900), 539.
Jeans supposes, in his 1905 paper "On the Partition of Energy between Matter and Aether",\(^{14}\) that a long time is required for the energy to be transferred to oscillation modes of high frequency.

Jeans takes the case of a system of gas and aether enclosed with perfectly reflecting walls, and expresses the energy of the system as the quadratic sum of the Lagrange coordinates and respective momenta of the system. If there are no interactions among the \(N\) molecules in the gas and the aether, the energy of the gas is the sum of the molecular kinetic energies, and the energy of the aether can be expressed in the form \(\Sigma a_i \xi_i^2\), using the Fourier components \(\xi_i\)'s of the electric and magnetic fields, some of which may be eliminated with the use of supplementary conditions, such as that for the electric field to be divergenceless, and so on. When the interactions among the molecules and aether are weak, even if they exist, the Lagrange coordinates and momenta remain approximately the same as those in the case of no interactions, and the equations of motion for them in the Hamilton form are still valid, providing the condition for realization of the equipartition of energy.

Jeans thinks, however, that "the partition of energy is not yet completely known" from this result only, since the degree of freedom of aether is infinite. He thus assumes that "the transfer of energy from matter to aether does not occur at all on the free-paths of molecules, but that vibrations are set up in the aether at collisions". Jeans supposes that the frequencies of the excited aether waves are smaller than or of the same order of magnitude as \(1/\tau\), \(\tau\) being the average collision time, and that the amplitudes of waves of frequencies much higher than \(1/\tau\) are very small.\(^{15}\) The transfer of energy is, therefore, considered to take place among the gas molecules and a finite number of the normal vibrations of the aether.

The frequency \(\nu\) of a normal oscillation of the aether in a cube of side length \(l\) is given by \(\nu = (c/2l)\sqrt{p^2 + q^2 + r^2}\), where \(p\pi/l\), \(q\pi/l\) and \(r\pi/l\) are the wave numbers in the directions of the \(x\), \(y\), and \(z\) coordinates, respectively, in terms of integers \(p\), \(q\) and \(r\). Because the number of the sets of \(p\), \(q\) and \(r\) satisfying the condition \(\sqrt{p^2 + q^2 + r^2} \leq \theta\) is approximately \(\pi\theta^3/6\), the number of the normal oscillations in the frequency interval \(d\nu\) turns out to be \((8\pi l^3/c^3)\nu^2 d\nu\). Distributing the energy \(kT\) equally to each normal oscillation, one gets for the

\(^{14}\)J. H. Jeans, Phil. Mag. 10 (1905), 91.

\(^{15}\)J. H. Jeans, Phil. Mag. 6 (1903), 279.
energy density of the radiation the formula

$$\rho_\nu d\nu = kT \frac{8\pi}{c^3} \nu^2 d\nu,$$  \hspace{1cm} (15)

which is Rayleigh–Jeans' distribution formula.

Expressing it as a function of the wave length $\lambda$, the total energy of the radiation can be given as

$$8\pi kT \int_{\lambda_0}^{\infty} \lambda^{-4} d\lambda + \int_0^{\lambda_0} f(\lambda, T, t) d\lambda.$$  \hspace{1cm} (16)

Here, the second integral expresses the energy of the aether vibrations of wave lengths shorter than $\lambda_0$. It depends on the time $t$ after the enclosure of the aether, because the transfer of energy to such degrees of freedom takes place slowly.

About the process in which energy is distributed from the heated gas to the aether, Jeans considers the following. At first, the energies of the degrees of freedom of the gas are distributed according to Maxwell's law. The time necessary to this is a small fraction of one second. After this, there arises an equalization of temperature in the gas by thermal conduction, and a transfer of energy between the principal degrees of the molecules and the vibrations of low frequencies in the aether. The time required for these phenomena must be measured in minutes, days, or centuries, according to the linear scale of the system. It is the transfer of energy to the vibrations of short wave lengths in the aether that comes lastly. If the gas is not too hot, a time of millions or billions of years must be needed for this phenomenon to take place.

Jeans stresses this point also in a 1908 paper "On the Theory of Radiation", and tries to "defend" Rayleigh–Jeans' distribution formula. When a hot body is enclosed in a cavity with perfectly reflecting walls, the body and the aether arrive finally at a state in which the law of the equipartition of energy is satisfied. For the system to be brought into this state, however, the system must not be radiative of energy, for a time duration measured in centuries or in ages. If the spectrum of the radiation in the cavity is observed through a small opening on the wall, the energy radiated from the hot body is maintained only for a time interval comparable with a thousandth of second, even under such a favorable condition for a spherical cavity of diameter of thirty centimeters, that the area of opening is a millionth of that of the whole.

---

wall. The energy which is analyzed in the spectroscope is, as long as short waves are concerned at least, to be considered as the energy coming directly from the body. There is no time for the energy to obey the law of equipartition. Considering this, Planck’s distribution formula with a maximum in the spectrum and Rayleigh–Jeans’ one without any maximum are identical practically, and are not distinguishable experimentally. This is the point of Jeans’ opinion to defend Rayleigh–Jeans’ distribution formula.

It makes an interesting contrast that on one side Jeans derived Rayleigh–Jeans’ distribution formula, which is valid in the region of long wave lengths and obeys the law of the equipartition of energy, on the basis of a fixed average collision time \( \tau \), and on the other side Wien found out his distribution formula, which is valid in the region of short wave lengths and does not obey the law of equipartition of energy, on the basis of Michelson’s variable time of passage \( \tau = \text{const.} / \nu \).

§3. Difficulties about Planck’s complexion

Equation (13), which is assumed by Planck to derive his distribution formula, gives for the entropy per resonator

\[
S_\nu = k \left\{ \left(1 + \frac{U_\nu}{h \nu} \right) \log \left(1 + \frac{U_\nu}{h \nu} \right) - \frac{U_\nu}{h \nu} \log \frac{U_\nu}{h \nu} \right\}. \tag{17}
\]

Planck interpreted this as the entropy due to the possibility of the resonator of taking exclusively energies which are integer multiples of an energy element \( \varepsilon = h \nu \).\(^{17}\) That is, Planck assumed for the number of ways of distributing \( P \) energy elements to \( N \) resonators the “complexion”

\[
W = \frac{(N + P - 1)!}{(N - 1)!P!}. \tag{18}
\]

Putting the entropy \( S_N \) of the system of \( N \) resonators to be \( k \log W \) after Boltzmann, and making use of the relations \( N! \approx N^N \) and \( U_N = N U_\nu = P \varepsilon \), Planck obtained the entropy per resonator \( S_\nu = S_N / N \), which is given by Eq. (17). As is noted on p. 27 of Vol. I, Boltzmann used an energy element \( \varepsilon \) as a mean for treating a continuous energy distribution and took the limit \( \varepsilon \to 0 \).

Einstein analyzed, in a 1909 paper "On the Present State of the Problem of Radiation,"\(^{18}\) the difficulty inherent in Planck's complexion. "To my thinking it is not to be doubted that our present-day theoretical view arrives necessarily at the law advocated by Mr. Jeans," says Einstein in this paper. "We must ask ourselves, in what relation does Planck's radiation theory stand theory with our, staying on the theoretical foundations recognized at present? The answer to this question is, to my thinking, made difficult by the fact that certain logical incompleteness attaches to Planck's representation of his own theory". Mentioning so, Einstein "shortly" makes a study of this problem as follows.

If one takes the standpoint that the irreversibility of natural processes is only of appearance, and that the irreversible process consists in a transition to a more probable state, one must first give a definition to the probability of a state. The only such definition that can be taken into consideration is the following.

Let \(A_1, A_2, \ldots, A_l\) be all the states that can be taken by a closed system with a definite energy, that is, all the states that we can distinguish by some means. According to the classical theory the system takes one of these states (\(A_l\) for example) after some time, and remains afterward in this state (thermodynamical equilibrium). According to the statistical theory, however, the system takes all the states \(A_1, A_2, \ldots, A_l\) again and again in an irregular sequence. (That this understanding is supportable comes directly from the properties of Brownian motion.) The system holds the state \(A_\nu\) only for a part \(\bar{\tau}\) of a very long time \(\Theta\). We name the limiting value of \(\bar{\tau}/\Theta\) the probability \(W\) of the state \(A_\nu\).

From this definition of \(W\) one can show that there should hold the equation for the entropy \(S\)

\[
S = \frac{R}{N} \log W + \text{const.}
\]

in which the constant is the same for states of the same identical energy.\(^{19}\)

Neither Boltzmann nor Planck has given the definition of \(W\). They have purely formally put \(W = \) the number of the complexions of the state under consideration. Now, if one defines the probability of the complexion, analogously to that of the state, by demanding that these complexions be equally probable,


\(^{19}\)\(k = R/N\) is equal to the Boltzmann constant. As will be seen later, until the later half of the paper under consideration, Einstein does not use the Boltzmann constant and the Planck constant in Planck's distribution formula. He uses only those physical constants which are well founded. (\(R\) is the gas constant and \(N\) is the Avogadro number.)
one arrives just at the definition of the probability of the state mentioned above. The relation between $S$ and $W$ is valid only when the probability of the complexion is defined in this way.

From these analyses, Einstein concludes that “in the resonator theory of radiation Mr. Planck would not be free in the choice of the complexion”. “He could set up the pair of the equations

$$S = (R/N) \log W,$$

$$W = \text{the number of the complexions},$$

only when he had supplemented the condition that the complexions should be so chosen that they would be found with equal probability on the ground of statistical considerations in the picture chosen by him. He would in this way arrive at the formula being defended by Jeans. The more each physicist should be glad that Mr. Planck ignored this demand in a very lucky way, the less it would be adequate to forget that the Planck’s radiation formula is incompatible with the theoretical foundation from which Mr. Planck has started”.

Although Planck’s complexions included such difficulties as are described above, Debye derived Planck’s distribution formula in a paper in 1910 “The Concept of Probability in the Theory of Radiation,”\(^\text{20}\) by applying Planck’s complexion to the radiation, on the supposition that energy is distributed among the oscillation modes of the radiation, also in units of the energy quantum $h\nu$ as to the resonators. In this work, Debye first calculates the number of the normal oscillations of the radiation in the frequency interval $d\nu$,

$$N_\nu d\nu = \frac{8\pi l^3}{c^3} \nu^2 d\nu,$$  \hspace{1cm} (19)

$l^3$ being the volume of the cavity, with a method similar to that of Jeans, and calls it “the number of the elementary states”. Then if, to this number of the elementary states, $f(\nu)N_\nu d\nu$ energy quanta are distributed, the energy of the radiation $u_\nu d\nu$ is given by

$$u_\nu d\nu = h\nu f(\nu)N_\nu d\nu.$$  \hspace{1cm} (20)

For the number $w_\nu$ of the ways to distribute $f(\nu)N_\nu d\nu$ energy quanta to $N_\nu d\nu$ elementary states, Debye sets up the equation

$$w_\nu = \frac{(N_\nu d\nu + N_\nu f d\nu)!}{(N_\nu d\nu)!(N_\nu f d\nu)!},$$  \hspace{1cm} (21)

making use of Eq. (18) of Planck.\textsuperscript{21) From the total number of ways of distribution \( W = \Pi w_\nu \), the entropy of the whole radiation \( S \) becomes, through the relation \( S = k \log W \),

\[
S = k \sum_{\nu} N_\nu \{(1 + f) \log(1 + f) - f \log f\} d\nu .
\]

Looking for \( f(n) \) which makes \( S \) maximum under the subsidiary condition that the energy of the whole radiation \( U = \int u_\nu d\nu \) is constant, and determining the Lagrange multiplier concerned with the use of the relation \( (\partial S/\partial U) = (1/T) \), he finds

\[
f(\nu) = \frac{1}{e^{h\nu/kT} - 1},
\]

from which, together with Eq. (20), Planck's distribution formula is obtained for the energy density of radiation \( \rho_\nu = u_\nu/l^3 \).

In this derivation of Planck's distribution formula by Debye, thus "it is not necessary to enter into the characteristic of the resonator going over the quantum hypothesis," as he wrote himself. He wrote moreover that "the above consideration does not explain the question whether the existence of the elementary quantum should be seen as the characteristic of the aether or as the characteristic of the matter," and that in the former case the calculation of \( w_\nu \) "could no more be immediately guaranteed". In this way, Debye treated the normal oscillations of the radiation and made no mention of the theory of light quantum, because of his consideration that if changes in the energy of the resonator should take place with the elementary quantity \( h\nu \), those in the energy of the radiation should do alike.

As will be described later in Sec. 2, §4, Chap. 5 of the present volume, the derivation of Planck's distribution formula in Debey's work is mathematically the same as the formulation of the statistical mechanics, which Einstein reached in his completion of Bose's work in 1924 to derive Planck's formula from the standpoint of the light quantum theory (the Bose--Einstein statistics for the gas of light quanta). In the stage of 1910, however, the conflict between Rayleigh--Jeans' formula, which comes from the equipartition law of energy demanded by the statistical heat theory, and Wien's formula, which is born out of the atomic theoretical viewpoint and does not obey the equipartition law, was confined within the property of Planck's resonator or the character

\textsuperscript{21)} The number 1 in Eq. (18) is neglected here in comparison with \( N \).
of Planck's complexion, so that Planck's formula remained a formal joining of the conflict. To approach closely the essence of this conflict, "the knowledge of the substantialistic structure of the matter was not yet enough."

2. Molecular Theoretical Significance of the Planck Theory

§1. Atomic vibration model for specific heat of solid body

In the 1907 paper "The Planck Theory of Radiation and the Theory of Specific Heat" cited above, Einstein clarified "in what sense the molecular kinetic theory of heat must be modified, in order to be brought in harmony with the distribution law of the black body radiation". According to Einstein, it must be assumed in the Planck theory that "for oscillatory ions which can mediate energy exchange between matter and radiation, the manifold of states which they are able to take should be less than in the case of the bodies in our experience. We should in fact assume that the energy of the "elementary thing" (Elementargebild) can take exclusively the values 0, \((R/N)\beta\nu\), \(2(R/N)\beta\nu\), etc.

"If the Planck theory of radiation strikes the core of matters, we must expect to find also in other regions of the heat theory such contradictions between the present molecular kinetic theory and the experience, that have to be solved in the adopted way". Stating so with an accurate methodology, Einstein now discusses the specific heat of the solid body.

On the assumption that each atom in a solid body makes a sinusoidal oscillation around its equilibrium point, the specific heat of a solid body per mol becomes \(C = 3Rn\), where \(n\) denotes the number of atoms per molecule. It is known that this relation is well satisfied for most elements and for many chemical compounds (Doulong–Petit's law, Neumann–Kopp’s rule). However, the elements C, B and Si possess at a normal temperature specific heats much less than the value given by this relation, and all solid chemical compounds which contain O, H or at least one of the elements mentioned above have specific heats less than that value.

Further, Drude showed, from the optical phenomena of dispersion, that many independently movable "elementary masses" belong to each atom in

---

22) Cf. p. 33, Sec. 3, §1, Chap. 1 Vol. I.
23) Cf. Ref. 2), on p. 7, Sec. 1, §1 of the present chapter.
24) \(R/N = k, \beta = h/k\). Cf. Footnote 19) on p. 16 of the present chapter.
chemical compounds, and that infrared eigenfrequencies are attributed to oscillations of atoms, and ultraviolet eigenfrequencies to those of electrons. This results in a specific heat per mol greater than $3Rn$, because the number of movable mass points per molecule becomes greater than $n$.

However, Einstein consider that, if the carriers of heat in a solid body are supposed to be periodically oscillating things with an energy-independent frequency $\nu$, then according to the Planck theory of radiation their average energy should be taken like that given by Eq. (2). The specific heat of the solid per mol should thus be taken as

$$C = 3Rn \frac{e^{\beta \nu/T} (\beta \nu/T)^2}{(e^{\beta \nu/T} - 1)^2}.$$  \hspace{1cm} (24)

Though this gives a value which is almost not different from $3Rn$ when $T/\beta \nu > 0.9$, the elementary thing under consideration does not contribute markedly to the specific heat when $T/\beta \nu < 0.1$.

Since $T/\beta \nu < 0.1$ at $T = 300$ K corresponds to $\lambda < 4.8 \mu$ for the wave length of light $\lambda$, Einstein concludes that “those electrons supposed for the explanation of the ultraviolet eigenfrequencies could not contribute to the specific heat at normal temperatures”. He also concludes that “as the carriers in the solid body (isolator) only positive ions are to be considered,” since $T/\beta \nu > 1$ at $T = 300$ K corresponds to $\lambda > 48 \mu$, and for infrared eigenfrequencies in general, $\lambda > 48 \mu$. Concerning this point, Einstein writes, in an errata about three months later, that “from Drude’s study, the supposition is not approved that every oscillatory elementary thing, which behaves as the carrier of heat, possesses always an electric charge”.

Einstein closes the paper under consideration by analyzing the temperature dependence of the specific heat of diamond with the use of Eq. (24). Determining the value of $\nu$ from the experimental value of the specific heat at $T = 331.3$ K, he gets the corresponding wave length $\lambda(=c/\nu) = 11.0 \mu$. The values of the specific heat calculated with this value of $\nu$ agree almost well with the experimental values at eleven temperatures between $T = 222.4$ K and $T = 125.8$ K. Einstein concludes that “we have to take therefore that the carrier of heat in diamond is a nearly monochromatic thing”.

In this way, it was made clear by Einstein that the quantum nature of energy, which was introduced as an attribute of Planck’s abstract resonator, belongs also to solid bodies which can be grasped by our hands. However, the clarification of the substance of the oscillatory thing and of its frequency
remained to be solved. So did the clarification of the problem why electrons in a metal, as a kind of solid, do not contribute to its specific heat.

§2. **Theory of specific heat of solids in terms of the quantum theory of sound**

Einstein attempted in a 1911 paper "A Relation Between the Elastic Behavior and Specific Heat of Solid Bodies Consist of Monoatomic Molecules," to derive the frequency of the atoms that have to do with the specific heat of a solid from its compressibility. The motive of this study was, as he wrote in the paper, Sutherland's work which showed that infrared frequencies are of the same order of magnitude as those of the transverse elastic waves of long wave length in solid bodies.

Now, Einstein's research begins with the supposition that a solid body consisting of monoatomic molecules makes a cubic lattice, in which each molecule interacts only with its respective neighboring molecules. The strength of the repulsive force between a pair of nearby molecules when they get nearer to each other by a unit distance from their lattice points is assumed. On the approximation that the neighboring molecules are distributed uniformly on a sphere, he calculates from the assumed repulsive force the restoring force of a molecule for its displacement from its lattice point, to obtain the molecular eigenfrequency. Also from the repulsive force, he next calculates the energy change of the solid body when it is compressed uniformly. Equating this energy change to that given in terms of the relative ratio of the volume reduction and the compressibility, he expresses the compressibility in terms of the constants of the repulsive force. The molecular eigenfrequency is thereby given by the molecular weight, density and compressibility of the solid body.

Applying this result to Ag, Einstein gets for the wave length corresponding to the eigenfrequency the value $\lambda \cdot 10^4 = 73$ cm. He says that the agreement between this value and the value $\lambda \cdot 10^4 = 93$ cm, which is obtained by Nernst by a comparison of Eq. (24) with the experimental data of the specific heat of Ag, is "really surprising". Einstein then concludes that "a more exact examination

---

27) W. Sutherland, Phil. Mag. 20 (1910), 657.
28) Einstein added in an errata [Ann. d. Phys. 34 (1911) 590] that he overlooked Madelung's work [Zeit. Phys. 11 (1910), 898], which "noted at first this fundamental important relation".
of Sutherland's understanding may be achieved only by completing the theory of solid bodies”.

In another paper in 1911, “Elementary Considerations on the Thermal Molecular Motions,” Einstein calculates the energy of a monoatomic molecule which makes an unharmonic oscillation in a solid body, and shows that it has, on the time average, a change comparable with the oscillatory energy itself within a half of the period. This is due to the fact that the neighboring molecules are not at rest but are in oscillation too. Einstein considers therefore that the molecular frequency is not definite but distributes over a certain region.

According to Nernst and Lindemann, on the other hand, the equation which is derived from the average of two equations corresponding to Eq. (24) with frequencies $\nu$ and $\nu/2$, reproduces the experimental temperature dependence of the specific heat of solid bodies much better than Einstein’s Eq. (24). Einstein takes notice of their result, and thinks that this is easy to understand by considering that each molecule oscillates with the frequency $\nu$ for one half of the time and with the frequency $\nu/2$ for the other half.

Generalizing this idea, Einstein pointed out, in a note added in proof to this paper, that the specific heat of non-monochromatic things can be expressed in the form

$$\int_0^\infty \Phi \left( \frac{\nu_0}{T} \right) \phi(x) dx,$$

where $\Phi(\nu_0/T)$ is the specific heat of the monochromatic things with frequency $\nu_0$, and $\phi(x)$ with $x = \nu/\nu_0$ is a function of to be taken as the ratio of repetition of the temporary frequency $\nu$. Following Einstein’s thought that the frequency taking part in the specific heat of the solid is related to its compressibility and is distributed over a certain region, Debye treated with success the oscillations of the molecules in the solid as those of sound waves.

In other words, Debye, in his 1912 paper “On the Theory of Specific Heats,” considers to analyze the molecular oscillations into Fourier components and characterize with their frequencies those taking part in the specific heat. As Debye says in this paper, “this stands in a direct analogy to Jeans’ proof of Rayleigh’s radiation formula”.

Jeans discussed, in his 1908 paper cited previously, sound wave oscillations in a vessel as an analogy to aether oscillations in a cavity. When some energy is given to a number of bells in a vessel which does not pass through sound, their energies become distributed with time to sound wave oscillations. In the final states the energies are dispersed among the random molecular motions. It was the point of analogy taken by Jeans that, if the random molecular motions are expressed in terms of regular sound wave oscillations by a Fourier analysis, the energies of the sound wave oscillations of long wave lengths obey a distribution formula of the Rayleigh–Jeans type.

Now, Debye considers a solid body as consisting of \( N \) monoatomic molecules, which he regards as mass points. Because the total degree of freedom of a solid body is \( 3N \), the solid body has \( 3N \) normal frequencies. In order to obtain these frequencies from \( 3N \) equations of motion for \( N \) molecules, it is necessary to make some assumption about the inter-molecular force. Debye thinks, however, that for discussion of specific heat, it suffices to know the characteristic feature of the distribution of the frequencies. He thus calculates, approximating the solid body by a continuous medium, the acoustic spectrum of the solid with the use of the elastic equation. The acoustic spectrum so calculated agrees well with the actual one for frequencies not too high, but it would only be an approximate one in the case of wave lengths comparable with the inter-atomic distance, as Debye notes.

From the fundamental equation for elastic waves, rewritten in terms of the scalar and vector potentials, he derives the equation to determine the normal frequencies of a spherical solid body, from which the number \( z \) of the normal frequencies smaller than \( \nu \) is obtained with an asymptotic approximation for high values of \( \nu \) to be

\[
z = \nu^3 VF, \tag{26a}
\]

with

\[
F = \frac{4\pi}{3} \rho^{3/2} K^{3/2} \left[ 2 \left( \frac{2}{3} \frac{1 + \sigma}{2\sigma} \right)^{3/2} + \left( \frac{11 + \sigma}{3} \frac{1 - \sigma}{1 - \sigma} \right)^{3/2} \right]. \tag{26b}
\]

Here, \( V, \rho, \kappa \) and \( \sigma \) denote respectively the volume, density, compressibility and Poisson ratio of the solid body. By making use of the velocities of the

\[32^3\text{J. H. Jeans, Ref. 16) on p. 14, Sec. 1, §2 of the present chapter.}

\[33^3\text{Quantities of the relative order of } 1/N \text{ are neglected.}\]
transverse and longitudinal waves \( c_t \) and \( c_l \), respectively, Eq. (26) given by Debye can be expressed as

\[
z = \frac{4\pi}{3} \nu^3 V \left( \frac{2}{c_t^3} + \frac{1}{c_l^3} \right), \tag{27}
\]

from which the similarity to the case of the aether oscillation discussed by Jeans is seen clearly (the numerical factor 2 in the parentheses being the number of freedom of the transverse wave).

From the relation \( 3N = \nu_m^3 \, VF \) for the maximum value \( \nu_m \) of the allowed frequencies, the number of the normal oscillations of the sound wave in the frequency interval \( d\nu \) becomes \( dz = (9N/\nu_m^3) \nu^2 d\nu \). The application of Planck's formula to the distribution of the energy of the sound wave leads to the following equation for the energy \( U \) of the solid body at the temperature \( T \),

\[
U = \frac{9N}{\nu_m^3} \int_0^{\nu_m} \frac{h\nu}{e^{h\nu/kT} - 1} \nu^2 d\nu. \tag{28}
\]

The specific heat \( C \) of the solid body is given by \( C = \partial U/\partial T \). \( C \) can be expressed as a function of \( T/\theta \) with \( \theta = h\nu_m/k \). \( \theta \) is nowadays called the Debye temperature.

Debye compared the experimental data for the specific heat of Al at 10 temperatures between 32.4 K and 331 K with the theoretical expression, and obtained \( \theta = 396 \text{ K} \). This value was in good agreement with the value \( \theta = 399 \text{ K} \) calculated from the compressibility, density and Poisson ratio of Al. The pair of values of \( \theta \) were 309 K and 329 K for Cu, and 215 K and 212 K for Ag, respectively.

At nearly the same time as Debye, Born and v. Kármán calculated in their paper "On the Oscillation of Space-Lattice"\textsuperscript{34}) the specific heat of solid. However, as they stated in their second paper "On the Theory of Specific Heat,"\textsuperscript{35}) "the priority in the exact formulation and an approximate solution of the problem belongs to Mr. Debye with certain days". In the first paper, they discussed the distribution of the frequency of lattice oscillation by assuming that two mass points on a pair of neighboring lattice points exert on each other a force proportional to the change in the mutual separation. And, in the second paper, they related the frequency distribution to the elastic constants

\textsuperscript{34}) M. Born and Th. v. Kármán, Phys. Zeit. 13 (1912), 295.

Conflict between the Wave and Particle Natures

of the lattice oscillation and of the solid, showing the correspondence of their theory to Debye's.

In this way, it was made clear that Planck's distribution formula applies well not only to waves of the radiation (light waves), but also to waves in a solid body (sound waves). Although the understanding of Planck's distribution formula itself was not yet reached, the universality of Planck's phenomenological formula was advanced.

Debye applied the method, which Jeans used in his attempt to defend Rayleigh-Jeans' distribution formula by making use of the example of sound waves in a gas, to sound waves in a solid body by taking notice of Einstein's idea. This makes an interesting contrast in the development of the quantum theory.

3. Conflict between the Wave and Particle Natures

§1. Fluctuation of radiation energy

As is seen in Chap. 1 of the present volume, it was clear that the Planck theory possesses many difficulties, such as the contradiction with the law of equipartition of energy, the contradiction with the electromagnetic theory, and the formal joining of Wien's distribution formula supporting the theory of light quantum to Rayleigh-Jeans' formula based on the theory of electromagnetic waves. However, it still could not be said that the difficulties in the theory of radiation were fully revealed. Einstein thus showed definitely that the very core of the difficulties in the theory of radiation lies in the conflict between the wave nature and particle nature of radiation, by analyzing fluctuations of the radiation field.

In the 1909 paper "On the Present State of the Problem of Radiation" quoted before, Einstein considers that "because the modification mentioned just now of the foundation of the Planck theory leads to a thorough change of our physical theory, it is very important to look for possible, simple and mutually independent interpretations of Planck's radiation formula, or the law of radiation in general as long as this may be presumed to be known". So, he takes up the problem of the fluctuations of the energy and momentum of radiation, as the ones "outstanding for their simplicity".

36) A. Einstein, Ref. 18), on p. 16, Sec. 1, §3 of the present chapter.
According to him, the equation \( S = (R/N) \log W \) may be used to determine the probability of a state of a closed system from the entropy \( S \) determined with the help of experience, contrary to the calculation of \( S \) from \( W \) determined by some complete theory. Any theory which provides for the state probability a value other than the one so determined is clearly to be rejected. In his paper on the theory of light quantum (1905), Einstein studied the statistical characteristics of thermal radiation, being based on Wien’s distribution formula that is valid for limiting values of large \( \nu/T \). This time he makes a similar analysis of Planck’s distribution formula.

Suppose that there are two spaces of volumes \( V \) and \( v \) that are connected to each other and covered by perfectly reflecting walls. Thermal radiation of frequencies in the interval \( dv \) is assumed to be enclosed in these spaces. Let \( H \) and \( \eta \) be the momentary radiation energies in \( V \) and \( v \) respectively. After a certain time, there holds the proportional relation \( H_0/\eta_0 = V/v \) to a good approximation. \( \eta \) at an arbitrary time deviates from \( \eta_0 \) according to a statistical law

\[
dW = \text{const.} \times e^{(N/R)S} d\eta,
\]

which proves valid directly when the relation between \( S \) and \( W \) is transformed into a differential form.\(^\text{37}\) Denoting the entropies in \( V \) and \( v \) by \( \Sigma \) and \( \sigma \) respectively and putting \( \eta = \eta_0 + \varepsilon \), one has

\[
S = \Sigma + \sigma = \Sigma_0 + \sigma_0 + \left\{ \frac{d(\Sigma + \sigma)}{d\varepsilon} \right\}_0 \varepsilon + \frac{1}{2} \left\{ \frac{d^2(\Sigma + \sigma)}{d\varepsilon^2} \right\}_0 \varepsilon^2 + \ldots.
\]

If \( V \) is assumed very large compared with \( v \) one has

\[
\left\{ \frac{d(\Sigma + \sigma)}{d\varepsilon} \right\}_0 = 0,
\]

so that \( S \) is given by

\[
S = \text{const.} + \frac{1}{2} \left( \frac{d^2\sigma}{d\varepsilon^2} \right)_0 \varepsilon^2 + \ldots.
\]

\(^\text{37}\)Equation (29) is due to the fact that \( W \) is given by \( W = \text{const.} \times \exp[(N/R)S] \) from the relation \( S = (N/R) \log W + \text{const.} \), and is proportional to the number of states with energies in a sufficiently narrow range. According to Einstein’s expression in his paper on the critical fluorescence of liquid [Ann. d. Phys. 33 (1905), 1275], \( W \) is proportional to “the probability of the domain of the state,” so that the probability of a quantity to be in the interval between \( \alpha \) and \( \alpha + d\alpha \) is proportional to \( \exp[(N/R)]d\alpha \).
If one is satisfied with the first order term in this expansion, which gives a very small error when $v$ is large compared with $\lambda^{3,38}$, $\lambda$ being the wave length, one gets with the use of the relation $d\eta = d\varepsilon$

$$dW = \text{const.} \times \exp \left[ \frac{1}{2} \frac{N}{R} \left( \frac{d^2 \sigma}{d\varepsilon^2} \right)_0 \varepsilon^2 \right] d\varepsilon. \quad (31)$$

The mean square of the fluctuation of the radiation energy in $v$ therefore becomes

$$\overline{\varepsilon^2} = -1 \left/ \frac{N}{R} \left( \frac{d^2 \sigma}{d\varepsilon^2} \right)_0 \right.. \quad (32)$$

If the distribution formula of radiation is known, one can calculate $\sigma$ from it. Thus, regarding Planck's distribution formula as an expression of the experience, one has$^{39}$

$$\overline{\varepsilon^2} = \frac{N}{Rk} \left( h\nu\eta_0 + \frac{c^3}{8\pi\nu^2} \frac{\eta_0^2}{v} \right). \quad (33)$$

This equation can be written in the form

$$\overline{\varepsilon^2} = \left( h\nu\rho_{\nu} + \frac{c^3}{8\pi\nu^2} \rho_{\nu}^2 \right) v d\nu, \quad (34)$$

as is noted by Einstein in a footnote in his paper under consideration. (Notice the relation $\eta_0 = \nu\rho_{\nu} d\nu$ for the energy density $\rho_{\nu}$ of the radiation in the frequency interval $d\nu$.)

In this way, Einstein arrives at "an easily interpretable expression" for the fluctuation of the radiation energy found in a small volume. Now, he goes on to show that the usual theory of radiation cannot be united with this result.

According to the usual theory, the fluctuation of the radiation energy occurs because of the fact that the infinite number of rays confined in the space to constitute the radiation in $v$, interfere with each other to bring a momentary energy which is sometime greater and sometime smaller than the sum of the energies that the respective rays would bring if they do not interfere at all. Einstein studies this effect by using a dimensional analysis, instead of complex mathematics, in the following way.

$^{38}$\begin{equation}
\nu \gg \lambda^{3} \quad \text{is also the condition for} \quad H_0/\eta_0 = V/v \quad \text{to hold. Under this condition} \quad \sigma_0 \quad \text{may be considered constant.}
\end{equation}$

$^{39}$\begin{equation}
\text{Here Einstein expresses Planck's distribution formula, for the first time, with the explicit use of the Planck constant } h \text{ and the Boltzmann constant } k.
\end{equation}$
Assume that the following conditions are satisfied.

1. \( \bar{\varepsilon}^2 \) depends only on \( \lambda, d\lambda \) and \( \rho_\lambda \).

2. The ray energies in neighboring domains of wave length and volume are simply added independently, as long as these domains are large enough so that the fluctuation of the radiation energy is proportional to \( d\lambda \) and \( v \).

3. \( \bar{\varepsilon}^2 \) has the dimension of energy squared.

From these conditions one gets for \( \bar{\varepsilon}^2 \), except for a numerical factor of the order of 1, \( \rho_\lambda^2 \lambda^4 v d\lambda \) that corresponds to the second term in Eq. (33) or (34). When Eq. (32) is applied to Rayleigh–Jeans' distribution formula, instead of Planck's, it gives only this term.

If \( R/Nk \) is put equal to a constant of the order of 1, according to the determination of the elementary quantum by Planck,\(^{40}\) the first term in Eq. (33) or (34) is much larger than the second term, for the visible rays surrounding us everywhere. This is not in harmony with the usual theory.

As we have seen before, Einstein did not recognize any theoretical ground for the relation \( S = k \log W + \text{const.} \) (\( W \) being the number of the complexions), and considered Planck's distribution formula to be a kind of experimental formula. Up to this time, therefore, he used \( R/N \) and \( \beta \) respectively for the quantities \( k \) and \( h/k \) appearing in Planck's distribution formula. Now, as is pointed out by Einstein in a footnote to his paper under consideration, "one would certainly get \( R/Nk = 1 \) by accomplishing the consideration of interference suggested above". Planck's distribution formula contains Rayleigh–Jeans' and Wien's formulas as the limiting cases for long and short wave lengths respectively. In the expression for \( \bar{\varepsilon}^2 \) resulting from Planck's formula, the contributions from these two limiting cases appear as a simple sum, so that the characteristic of Planck's formula becomes much clearer. The equality of \( k \) and \( R/N \) used by Planck get rid of its contingency.

Going back to Einstein's analysis, if one puts \( R/Nk = 1 \) following Planck, the first term in Eq. (33) or (34) gives such fluctuation of the radiation energy as would arise if the radiation consists of independently moving point-like quanta of energy \( h\nu \). This can be seen by a simple calculation. The first term gives a large value for the relative fluctuation \( (\varepsilon^2/\eta_0^2)^{1/2} \) when \( \eta_0 \) becomes small,

\(^{40}\)This is to point out that Planck determined the constants \( k \) and \( h \) in his distribution formula from the experimental data on Stefan–Boltzmann's law and Wien's displacement law (cf. p. 32, Sec. 2. §9, Chap. 1, Vol. 1 of this book).
Conflict between the Wave and Particle Natures

and the relative fluctuation is independent of the size of \( v \). Einstein stresses that these two points show that “the actual statistical properties of radiation are fundamentally different from those that we should expect according to our usual theory, which is based on linear homogeneous differential equations”.

Let us now, for a moment, look back upon Planck’s calculation of the entropy of the resonator. After we have known Einstein’s analysis described above, one sees that Planck treated in effect a type of fluctuation about his resonator. As is shown in Sec. 1, §1 of present chapter, Planck calculated the change in the entropy \( dS_t \) of the system of a resonator and radiation after a short time \( dt \), when the energy of the resonator is changed from the stationary value \( U_\nu \), with the result given by Eq. (10) for the relation between \( U_\nu \) and the entropy per resonator \( S_\nu \). He then applied this result to \( n \) independent resonators and assumed Eq. (11), from the solution of which he got Eq. (12),

\[
\frac{d^2 S_\nu}{dU_\nu^2} = -\frac{\alpha}{U_\nu}.
\]

Integrating this equation, he arrived at Wien’s distribution formula.

Equation (11), which Planck considered to be self-evident, corresponds to the particle-nature of radiation that is clarified by Einstein with the first term of \( \varepsilon^2 \). In fact, Eq. (13)

\[
\frac{d^2 S_\nu}{dU_\nu^2} = -\frac{\gamma}{U_\nu(\beta + U_\nu)},
\]

which he assumed to get his distribution formula on the supposition that the procedure leading to Eq. (11) is “not easily understandable and anyway hardly provable,” becomes

\[
1 \left/ \frac{d^2 S_\nu}{dU_\nu^2} \right. = -\left( \frac{\beta}{\gamma} U_\nu + \frac{1}{\gamma} U_\nu^2 \right),
\]

when its reciprocal is taken, as was pointed out by Planck in his Nobel prize lecture (1918).\(^{41}\) This is nothing but the equation that gives the fluctuation of the energy of the resonator corresponding to Eq. (33) or (34).

§2. **Fluctuation of radiation pressure**

After the analysis of the fluctuation of the radiation energy, Einstein goes on to study the fluctuation of the radiation pressure. He considers the case of a mirror inside a cavity surrounded with a material of the temperature \( T \). The mirror is supposed to be freely movable in the direction of its normal. When the mirror moves with a velocity, because of the motion the radiation is reflected more on the front side than on the back side.\(^{42}\) The radiation pressure acting on the front side is therefore greater than that acting on the back side. Thus, because of the motion of the mirror relative to the cavity radiation, a force similar to a frictional force acts on the mirror, so that the momentum of the mirror should gradually diminish, if there is nothing to compensate on average the loss in the momentum of the mirror due to this force. On the other hand, corresponding to the irregular fluctuation of the radiation energy studied above, there is an irregular fluctuation of the radiation momentum, or that of the pressure exerted on the mirror by the radiation. As a result of the pressure of this kind, the mirror would go into motion if it is initially at rest. Einstein points out that the resulting motion of the mirror is very much analogous to the Brownian motion of a colloidal particle.

Let \( u \) be the velocity of the mirror at a time \( t \). If the mass of the mirror is \( m \) and the resistive force per unit velocity is \( P \), the velocity of the mirror diminishes by \( P ur / m \) in a short time \( \tau \). There is further more a change of velocity in the time \( \tau \) that results from the irregular fluctuation of the radiation pressure. Denoting it as \( \Delta \), the velocity of the mirror at the time \( t + \tau \) is

\[
 u - P r u / m + \Delta .
\]

For the velocity of the mirror not to change on the average in the time \( \tau \), we must have

\[
 \overline{(u - P r u / m + \Delta)^2} = \overline{u^2} .
\]

Neglecting small quantities of higher orders and noting that \( \overline{u \Delta} \) vanishes, this becomes

\[
 \overline{\Delta^2} = (2 P r / m) \overline{u^2} .
\]

\(^{42}\)In this sense, the mirror should be double-sided, though Einstein does not mention it explicitly.
\( u^2 \) in this equation can be rewritten by making use of

\[
\left( \frac{1}{2} \right) m u^2 = \left( \frac{1}{2} \right) RT/N,
\]

which is deduced from \( S = (R/N) \log W \).

The mirror is assumed to reflect completely the radiation of frequencies in the specified range from \( \nu \) to \( \nu + d\nu \), and allow to pass through completely the radiation of other frequencies. From a purely electrodynamical calculation, \( P \) turns out to be given by

\[
P = \frac{3}{2c} \left( \rho_\nu - \frac{1}{3} \nu \frac{d\rho_\nu}{d\nu} \right) d\nu f,
\]

(36)

\( f \) being the area of the mirror.\(^{43} \) One has therefore

\[
\frac{\Delta^2}{\tau} = \frac{3}{2} \frac{RT}{N} \left( \rho_\nu - \frac{1}{3} \nu \frac{d\rho_\nu}{d\nu} \right) d\nu f.
\]

(37)

This is transformed into

\[
\frac{\Delta^2}{\tau} = \frac{1}{c} \left( h\nu \rho_\nu + \frac{c^3 \rho_\nu^2}{8\pi \nu^2} \right) d\nu f,
\]

(38)

by making use of Planck’s distribution formula.

On this result Einstein states as follows. “From the usual theory the expression of the second term should again be deduced (fluctuation due to interference). If the first term only were retained, the fluctuation of the radiation pressure would be explained completely by the assumption that the radiation consists of independently moving, little spread complexes of energy \( h\nu \). The formula here also means, that according to Planck’s formula the effects of both the mentioned causes of fluctuation act like the fluctuations (errors) that arise from mutually independent causes (additive combination of the terms, of which the squared fluctuation is composed).”

Thus, Einstein concluded from these analyses of the fluctuation of radiation energy and that of radiation pressure, that “it turns out undeniable that the constitution of the radiation must be other than what we think at present,” and “it does not suffice to assume that the radiation could be emitted and

\(^{43}\)Einstein does not show the calculation of \( P \) “to simplify the description”. He gives a detailed calculation however is his 1916 paper “On the Quantum Theory of Radiation,” which will be considered later in Chap. 3.
Difficulties in Radiation Theory

absorbed only in quanta of energy $h\nu$, or that a property of emitting or absorbing matter would only come into question". These words of Einstein are important, because they show that in this paper he recognized Planck's distribution formula epistemologically, to go on to the clarification of its essence. Moreover, he added carefully that "it is to be stressed that the considerations given above would never lose its value essentially, even when Planck's formula turns out invalid; the part well confirmed by experience of Planck's formula (Wien's radiation law valid in the limit of large $\nu/T$) is the one that leads to the light quantum theory". Einstein emphasized in this way the necessity of innovating the theory of radiation.

§3. Conflict between the concepts of light wave and of light quantum

The fact that both the squared-means of fluctuation of radiation energy and of radiation pressure had such a duality as described above was to imply a serious conflict between the two concepts of light wave and light quantum. Einstein discussed this conflict in his 1909 paper "On the Development of Our View of the Essence and Constitution of Radiation".44) This is just the paper that we have planned in Sec. 3, §11, Chap. 1, Vol. I to treat in Vol. II, as one that presented the problem of unifying the wave and particle natures for the first time.

His discussions in this paper start from a summary of the history of how the change in our cognition of aether took place. The summary ends with the problem of the equivalence of radiation energy and mass, which we have treated in Sec. 4, Chap. 1, Vol. I. According to Einstein, "the theory of relativity has thus changed our view of the nature of light, as far as it does not grasp light as the result of the state of a hypothetical medium, but as something of independent existence like matter".

Einstein considers, however, "the theory of relativity changed nothing about our grasp of the structure of radiation, especially of the distribution of the energy in a space full of radiation". His sentence quoted here is a natural one, because while Wien’s distribution formula leads to the theory of light quantum, by which the theory of relativity was brought about, as we have emphasized in Vol. I, the peculiarity of Planck’s distribution formula remains to be solved. Einstein thinks that "our theory of light can not clarify certain

Conflict between the Wave and Particle Natures

characters of the phenomena of light,” and to such problems “the wave theory in its present state of understanding gives no answer”.

As an example, he takes the case of the production of cathode rays by X-rays. Primary cathode rays striking on a metallic plate $P_1$ produce X-rays, which then strike another metallic plate $P_2$ to produce secondary cathode rays. The velocities of the secondary cathode rays are of the same order of magnitude as that of the primary ones, and depend neither on the distance between $P_1$ and $P_2$ nor on the intensity of the primary cathode rays, depending only on the velocity of the primary cathode rays. If the intensity of the primary cathode rays or the area of $P_1$ is diminished, so that the collision of each electron of the primary cathode rays can be understood as an isolated process, there would be produced at $P_2$ either nothing, or an electron with energy of the same order of magnitude as that of the electron striking $P_1$. In other words, “the elementary radiation process seems to occur so, that the energy of the primary electron is not divided and dispersed through a spherical wave propagating into all directions as the wave theory demands,” and “the elementary process of emission of radiation seems to be directional”.

Einstein thinks that “the constitution of radiation seems, therefore, to be something other than the one that the wave theory makes us infer”. He then takes the Planck theory of radiation and summarizes the analysis he made of it from the view point of the statistical heat theory, which we have described in Sec. 1, §3 of the present chapter, to conclude that “to assume the Planck theory means, according to my thinking, just to reject the foundation of our theory of radiation”. He further takes the fluctuation of the radiation pressure described above, and notes the fact that $\Delta^2$ is given by the sum of two independent contributions due to light quanta and light waves. When the density of the radiation energy is low, the fluctuation due to light quanta is larger than that due to light waves, the square of the former being $6.5 \times 10^7$ times of that of the latter for $\lambda = 0.5 \mu$ and $T = 1,700$ K, as he points out.

Einstein concludes his discussions by saying “As far as I know, the establishment of a mathematical theory of radiation is not yet done well, that is adequate both for the wave structure and the structure following the first term of the above equation (quantum structure).... Let us imagine that the diffraction and interference phenomena were yet unknown, but we knew that the average value of the irregular fluctuation of the radiation density were determined by the second term of the above formula, in which $\nu$ is a parameter with unknown meaning to determine the color. Who would
have sufficient fantasy to construct the wave theory of radiation on this ground?"

Being based on the light quantum theory, Einstein thus emphasizes the necessity of recovering the wave nature of radiation that is confirmed by Planck’s distribution formula. In this connection, he adds the possibility, as an image for the time being, that the total energy of the electromagnetic field is localized at its singular points, as in the old theory of action at a distance. In this possibility, each singular point is supposed to be surrounded by “a force field, which possesses essentially the character of a plane wave, and whose amplitude diminishes with the distance from the singular point”. For many singular points in a distance small compared to the size of the force field of a singular point, their force fields are supposed to be superposed, so that there would arise in their totality a wave field of force, that would probably be little distinguished from a wave field in the sense of the usual wave theory of light.

His expression “a force field which possesses essentially the character of a plane wave” reminds us of his discussion about the directionality of radiation mentioned above. Although “no value is ascribed to such an image as long as it does not lead to an exact theory,” he wished to “exemplify that both the structural properties (wave structure and quantum structure), which should be attributed to the radiation because of Planck’s formula, are not to be seen as incompatible”.

Thus, it was made clear that the difficulty in the theory of radiation includes the conflict between the wave and particle natures of light. With the theory of light quantum, the hypothetical substance named aether was negated to be dissolved into a completely functional concept, and the light quantum entered as the new substance. This entirely new substance demands a new logic, so that it is to be grasped only in an essentialistic theory that unifies the functions of the particle nature, on which the light quantum itself stands, and of the wave nature, which the aether carried out.

Until quantum mechanics was established so that the wave and particle natures were grasped in a unified way, the conflict between the two natures caused various paradoxes. Lorentz urged in a paper in 1910, “The Hypothesis of the Light Quantum,” that the light quantum, such as described in Einstein’s

following sentence in his original paper on the light quantum (1905), could not exist from the viewpoint of the phenomena of interference and diffraction — "In the spreading of light rays outgoing from a point, the energy is not distributed over larger and larger growing space, but composed of a finite number of energy quanta localizing in space points, move without dividing into parts and can be absorbed and generated only as wholes.\textsuperscript{47)}

Following Lorentz, it is known that with homogeneous light one sees interference for a phase difference of even two million periods. The light under consideration should therefore proceed regularly over at least two million wavelengths (and probably not shorter than three million). He considers that this holds for the respective light quanta, because it could not be the case that the oscillations of two different light quanta are in phase or in opposite phase, and strengthen or weaken each other to form clear interference stripes. Thus, if the wavelength of the light is $4 \times 10^{-5}$ cm, each light quantum should spread over at least 80 cm in the direction of its motion.

Lorentz then argues that a peculiar difficulty would arise in the case of light absorption by a molecule, if each light quantum is not confined within a very small domain but ranges over many series of wave lengths. A molecule should be considered to absorb light by catching the energy of a large number of oscillations. One should therefore imagine that a molecule receives temporally the energy to catch it if the molecule gathers a whole quantum, or releases it if the molecule does not succeed in doing so.

Lorentz argues that a light quantum should spread also in directions transverse to its motion, from the viewpoint of the phenomenon of diffraction. Let us assume that a star is observed with a telescope, and the image is as good as is expected from the experimentally well confirmed theory of diffraction. The image becomes worse when the objective lens of opening area $O$ is stopped down to have an opening area $o$, the area of the image on the focus plane increasing from $s$ to $S$. This could not be the case if the light quantum is point-like and does not have any transverse spreading. Because between two different light quanta there is no relation, the light quanta passing through the full opening $O$ come into $s$, and this situation should not change when the objective lens is covered with a diaphragm to stop down the opening from $O$ to $o$. If the light quantum has certain transverse spreading, the image could change according to the diaphragm, as the light quanta entering within the

\textsuperscript{47)} Cf. p. 35, Sec. 3. §2, Chap. 1, Vol. I of the present book.
diaphragm would participate in forming the image. When the size of the diaphragm becomes smaller than the spreading of the light quantum, no light would be seen any more.

He then argues that the light quantum should spread over a domain almost comparable to the objective lens, by considering the case in which the left- and right-hand sides of the objective lens are covered alternately. This is because the image changes greatly each time of doing so. Accordingly, the spreading of the light quantum should be of an area of the order of several hundred cm$^2$ when the diameter of the objective lens is 50 cm, and of an order much greater than this for the "new" reflecting telescope of diameter 150 cm at Hale. Any light quantum coming from a star should be expected to have an exceedingly large spreading, since it would be accidental that the spreading of the light quantum coincides with the size of an optical apparatus one uses.

He argues further that, when one observes a star with his naked eyes, each light quantum could enter into his pupil only if it separates into fragments as small as one ten thousandth of the original light quantum. Many such fragments should be gathered up again in some way into a light quantum on the retina, because the light could act only as a whole quantum on the retina. In this way, Lorentz arrives at the conclusion that "it can not be possible for light quanta to be concentrated within small spaces and not divided always".
Chapter 2

The Quantum of Action and Atomic Models
Planck (1910)
Introduction of \( h \). Hypothesis of the quantum nature of energy.

Thomson (1903)
Model of atom without nucleus

Nagaoka (1903)
Model of atom with nucleus

Haas (1910)
Supposition that from the atomic radius assumed in Thomson's model and \( m, e \) of the electron, \( h \) would be derived.

Einstein (1909)
Opinion that a theory to give the constitution of the electron would give \( h = e^2 c \).

Rydberg (1890)
Conjecture of the combination rule

Ritz (1903)
Special mathematical models for the combination rule.

Ritz (1908)
Interpretation of the combination rule by means of a model representing atom as a series of bar magnets.

Establishment of the combination rule by systmatization of data, and prediction of the Paschen series.
Rutherford (1911)
Experimental verification of the model of atom with nucleus

Bohr (1912)
Determination of the number of electrons in a hydrogen atom; discussion of the relation of the characteristic X-ray to the frequency of an inneratomic electron.

Sommerfeld (1911)
The action integral $= h$.
Assertion that the existence of the atom would be due to $h$.

Planck (1911)
The phase integral $= h$.
Derivation of the quantum nature of energy.

Nicholson (1912)
Theory of coronal spectra on the basis of an atomic model of Nagaoka type, with indication of the quantum nature of angular momentum.

Bohr (1913)
The Bohr model of atom. Derivation of energy levels from the quantum nature of angular momentum. Grasping of the combination rule by the frequency relation.

Moseley (1913)
Verification of the Bohr model by X-ray experiment. Determination of atomic numbers.

Notes: 1) The numbers above a frame indicate Sec. and Subsec. 2) $e$, $c$, $m$ and $h$ are the elementary electric charge, light velocity, electron mass and the Planck const., respectively.
Chapter 2

The Quantum of Action and Atomic Models

Theme in Chapter

Einstein consider that in a new electromagnetic theory that would settle the problem of the self-energy of electron, the Planck constant \( h \) would be given in terms of the elementary electric charge and light velocity. In contrast to this, Haas supposed that \( h \) could be deduced from the atomic radius assumed in the Thomson model of the atom and the charge and mass of electron.

In connection with this, it should be greatly appreciated concerning in the history of the formation of quantum mechanics that Sommerfeld gave the view that the existence of the molecule is the function and result of the existence of \( h \).

It should never be forgotten also that, different from Ritz who tried to explain the combination rule of spectral lines with special mathematical models, Nicholson followed Sommerfeld’s view to discuss the spectra of coronae of the sun on the basis of an atomic model of the Nagaoka type with several electrons, in which no assumption is made of the atomic radius, and came to point out the existence of the quantum nature of angular momentum due to \( h \), together with an estimation of the atomic radius.

Bohr, who determined the number of electrons in the hydrogen atom from his analysis of the energy loss of \( \alpha \)-rays through hydrogen gas, and thus understood that the problem of the stability of atoms could not be solved with the classical mechanics, presented the Bohr model of the atom, in which the existence of atomic discrete energy levels results from the quantum nature of angular momentum, and the combination rule of spectral lines is grasped with the frequency relation.

1. Theoretical Significance of the Quantum of Action

\( \S 1. \) The quantum of action and fundamental physical constants

As is described in the previous chapter, the character of the Planck theory of radiation was gradually made clear by Einstein in a series of analyses beginning with his paper on the light quantum in 1905. Together with this, as regards
the new physical constant that the Planck theory brought in, that is, the constant that was named the quantum of action because of its dimension, and is nowadays called the Planck constant, discussions were made on the question how to relate it to other physical constants, or how to place it in the system of physical theories.

Einstein took notice of the fact that \( h \) has the same dimension as \( e^2/c \) (\( e \) and \( c \) being respectively the elementary electric charge and light velocity), and considered that it might be possible to deduce \( h \) from the relation \( h \sim e^2/c \) if the electromagnetic theory could be modified so that it would be able to deduce \( e \).

Let us now follow his discussions in the last part of the 1909 paper “On the Present State of the Problem of Radiation”,\(^1\) which was previously taken up to in the last chapter. After the recognition of Planck's distribution formula because of the result obtained for the fluctuation of radiation energy and radiation pressure, Einstein thinks that “from the fact that Jeans’ law appears valid in the limit (for small \( \nu/T \)), it seems to result that not a complete abandonment but only a modification of our present theory comes into question”. As an indication about how such a modification should be accomplished, he takes a dimensional analysis in consideration of that used by Jeans some years before, and tries to relate \( h \) to the fundamental physical constants.

Suppose that there is a space in which an ideal gas, ions and radiation are confined, and interchanges of energy take place between the gas and radiation on account of the charges of the ions. Einstein supposes that the following quantities are concerned in the determination of the energy density \( \rho \) of the radiation:

- (a) the average energy \( \eta \) of a molecular object (equal to \( RT/N \) except for an unspecified numerical factor),
- (b) the light velocity \( c \),
- (c) the elementary electric charge \( e \),
- (d) the frequency \( \nu \).

Though in Einstein’s paper no reference is given to Jeans’ paper, Jeans had made a dimensional analysis in 1905 in a paper “On the Laws of Radiation”\(^2\) to give the expression \( \rho = \lambda^{-4}kTf(c_1, c_2) \). In this study, Jeans considered that thermal radiation from a solid is due to motions of electrons, and supposed that

---

\(^1\)A. Einstein, Ref. 18), p. 16, Sec. 3, §1, Chap. 1 of the present volume.
the distribution law of radiation is determined by the mass of electron \( m \) and “the dielectric constant of aether” \( K \), besides the physical quantities selected by Einstein. In Jeans’ expression for \( \rho \), both the constants \( c_1 = kTm^{-1} \) and \( c_2 = \lambda kTke^{-2} \) are dimensionless and \( k = R/N \). He Neglected \( c_1 \) because of its small value \( 3.6 \times 10^{-8} \) and got \( \rho = \lambda^{-4} kTf(\lambda T) \), leaving out explicitly the universal constants \( k, K \) and \( e \) in \( c_2 \). Jeans supposed further that the function \( f \) would diminish as \( \exp(-\tau c/\lambda) \) for short wave lengths, where \( \tau \) denotes the average collision time.

Now, Einstein derives from his dimensional analysis

\[
\rho = \frac{e^2}{c^4} \nu^3 \psi(\alpha); \quad \alpha = \frac{Ne^2 \nu}{Rc T},
\]

where \( \psi(\alpha) \) is an unknown function. This satisfies Wien’s displacement law. Comparing this with Planck’s distribution formula, he gets

\[
\frac{h}{c^3} = \frac{e^2}{c^4} \quad \text{and} \quad \frac{h}{k} = \frac{N e^2}{R c},
\]

so that he has

\[
h = \frac{e^2}{c} \quad \text{and} \quad k = \frac{R}{N}.
\]

Although the actual values of \( h \) and \( e^2/c \) are different by three decimals, Einstein considers that this may be attributed to some unknown dimensionless numerical factors.

He emphasizes that “the importance of this derivation consists in that the light quantum constant \( h \) is reduced to the elementary quantum \( e \) of electricity”. The elementary charge \( e \) is, however, “a foreign thing in the Maxwell–Lorentz electrodynamics”, and it is necessary to introduce a force of a different kind to constitute the electron. Einstein thus states that “one usually introduces a strong frame which should prevent the electric mass of the electron from breaking up, under the influence of its electric interaction”. To him “the relation \( h = e^2/c \) seems to show that the same modification which includes the elementary quantum \( e \) will include also the quantum structure of the radiation as a result”. In such a modification, “the fundamental equation of the optics

\[
D(\varphi) = \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} - \left( \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} \right) = 0
\]
will be replaced by an equation in which the universal constant $e$ appears in a coefficient. The desired equation must, adds Einstein, lead to the form $D(\varphi) = 0$ in the limit of large amplitudes, at least in the case where Jeans' law is valid in the limit of small $\nu/T$.

Einstein aimed, in this way, to understand the quantum of action in connection with the elementary electric charge by finding out a new theory, although the quantum theory has not afterward developed in the direction that he thought of at that time. Still more, the problem of the electromagnetic mass of electron has brought into the quantum electrodynamics a serious difficulty, which is solved for the time being with the theory of renormalization after the theory of $C$-meson, but is not essentially understood yet.\(^3\)

§2. The Thomson model of the atom and the quantum of action

Haas took a standpoint, contrary to Einstein, within the framework of the usual theory, and attempted to deduce $\hbar$ under peculiar assumptions, from the radius of hydrogen atom and the elementary electric charge, by making use of the Thomson model of atom. He used inversely the relation so obtained for these quantities to evaluate numerically the radius of hydrogen atom and the elementary electric charge, from $\hbar$ and the Rydberg constant.

In a paper "On the Electrodynamical Meaning of Planck's Radiation Law and on a new Determination of Electric Elementary Quantum and the Dimension of Hydrogen Atom",\(^4\) Haas considers "the Thomson model of hydrogen atom as a special case of optical resonator". According to him, "in the case of the hydrogen atom, which includes only one electron and represents the simplest form of atom, the positive charge is consequently equal to the elementary electric charge".

By the way, as we have mentioned in Sec. 6, §1, Chap. 2, Vol. I (p. 182), "about 1905 various experiments were done to determine the number of atomic electrons. In particular, Thomson paid attention to the problem of the number of atomic electrons, and attempted to clarify the problem of atomic structure by synthesizing the result of those experiments. Thomson studied this problem

\(^3\)In the classical theory various attempts were done, about the difficulty that if the electron has no size its self energy becomes infinity, and if it has a size its electric charge breaks up as Einstein noted.

in 1906. In the next year, he also treated this problem in a book, and in another book too. We have mentioned also in Vol. I (p. 183), that Thomson investigated the intensity and absorption of cathode rays and dispersion of light, with the following conclusion — "The data at present available indicate that the number of electrons in the atom is equal to the atomic weight, but as the evidence is rather indirect and the data are not very numerous, further investigation is necessary before we can be sure of this equality."

Now, Haas cites the German edition in 1909 of Thomson's book mentioned above. Haas followed perhaps Thomson's conditional conclusion that the number of atomic electrons might equal to the atomic weight. As has been stated also in Sec.'s 4 and 5, Chap. 2, Vol. I, the number of electrons in the hydrogen atom was confirmed to be equal to one by Bohr in 1913 from the study of energy loss of charged particles through hydrogen gas, after the experimental confirmation by Rutherford in 1911 of the Nagaoka model of the atom with nucleus.

Using the radius of hydrogen atom in the Thomson model \( a \), that of the orbit of the electron \( r \), and the elementary electric charge \( e \), Haas concludes that the energy \( E \) of the electron in the hydrogen atom is given by \( \frac{e^2 r^3}{a^3} \) for \( r \leq a \), and by \( \frac{e^2 a^3}{r^3} \) for \( r > a \), so that it takes the maximum value \( \frac{e^2}{a} \) at \( r = a \). Haas then assumes that this maximum value is equal to \( h \nu \). Haas's expression of \( E \) for \( r > a \) is not right as he mistakes in the calculation of the potential energy of the electron for \( r > a \) by using \( \frac{e^2}{r^2} \times (r/2) \). Though \( E \) does not take the maximum value \( \frac{e^2}{a} \), if the electron is assumed not to move only within the sphere of the positive charge, it is important to Haas to

---

5) J. J. Thomson, Phil. Mag. 9 (1906), 769.
8) Haas calculates the potential energy for \( r \leq a \) elementarily by putting \( (e^2 r/a^3) \times (r/2) \), and says that for \( r > a \) it is "similarly" given by \( e^2/2r \). In the former case, where the force is proportional to \( r \), the potential energy is obtained by multiplying the force by the average displacement \( r/2 \), but in the latter case one has to integrate the force with respect to \( r \).

In a lecture at Solvay Conference (1911), which will be discussed in the next subsection §2, Planck introduces Haas's work, saying "In the case of the amplitude (of the electron) larger than the radius of the atom, the oscillation is impossible because the electron always goes away from the sphere." Planck may not have noted Haas's mistake, or may not have assumed further that electrons in a stable atom remain in the atomic sphere. Schidlof tried to modify Haas's calculation, in a paper [Ann. d. Phys. 35 (1911), 537] cited also in Planck's lecture, with the assumption that almost all the electrons, except those in the outmost ring and the valence electrons in a heavy atom, gather to form a sphere inside the sphere of positive charge, so as to restrict the domain only to \( r \leq a \).
assume the relation
\[ h\nu = \frac{e^2}{a}. \] (4)

From this, together with the frequency of the circular motion of the electron
\[ \nu = \frac{e}{2\pi a\sqrt{am}}, \] (5)

he gets
\[ h = 2\pi e\sqrt{am}, \] (6)

where \( m \) denotes the mass of the electron.

The assumption given by Eq. (4) to suppose \( e^2/a \) as the maximum energy of the electron is contrary to the Planck theory of radiation, in which the energy of the resonator is supposed to take integer multiples of \( h\nu \). To avoid this contradiction, Haas makes, further, a very peculiar assumption that the temperature which the atom can take has the maximum value \( T' \), where the average energy \( U_\nu \) of Planck's resonator becomes equal to \( h\nu \). That is, he puts
\[ \frac{h\nu}{\exp(h\nu/kT') - 1} = h\nu, \]
so that he gets \( \exp(h\nu/kT') = 2 \).

For the value of \( a \), he takes \( 2^{1/3} \) times the value \( 2.2 \times 10^{-8} \) cm that was given by Loschmidt for the radius of hydrogen molecule
\[ a = 5.6 \times 10^{-8} \text{ cm}, \]

and for the value of the specific mass of electron he takes
\[ e/m = 5.6 \times 10^{17} \text{ g}^{-1/2} \text{ cm}^{3/2} \text{s}^{-1/2}. \]

Haas says that for \( e \) the values of 3.1, 3.4 and \( 4.6 \times 10^{-10} \) c.g.s. are reported, from each of which he obtains for \( h \) the values of 6.1, 7.0 and \( 10.9 \times 10^{-27} \) c.g.s., respectively, in comparison with the value \( 6.548 \times 10^{-27} \) c.g.s. due to Planck.

Since the values of \( a \) and \( e \) are uncertain, Haas calculates them from that of \( h \), by using inversely Eqs. (4) and (5). For this purpose, it is necessary to know the value of \( \nu \). He looks for it in Balmer's spectral formula
\[ \nu = cN \left( \frac{1}{2^2} - \frac{1}{n^2} \right), \] (7)

\(^9\)In another paper [Phys. Zeit. 11 (1910), 537], Haas shows this equation without any comment about the large- or small-ness of \( r \) relative to \( a \).
The Quantum of Action and Atomic Models

where $N$ is the Rydberg constant ($109677.691 \text{ cm}^{-1}$)$^{10}$ That is, Haas assumes that the maximum wave number $\nu_{\infty} = cN/4$ given by Balmer's formula corresponds to his maximum energy $e^2/a$, and puts instead of Eqs. (4) and (5)

$$h\nu_{\infty} = \frac{e^2}{a},$$

and

$$\nu_{\infty} = \frac{e}{2\pi a\sqrt{am}}.$$  \hfill (9)

Using these formulae, he obtains $e = 3.18 \times 10^{-10}$ c.g.s. and $a = 1.88 \times 10^{-8}$ cm from the values of $h$ and $\nu_{\infty}$ given above.

As is described above, Haas's work contained a very peculiar assumption and a mistake. Equation (8), which resulted as an byproduct of his work, however, presented the possibility that the atomic radius could be deduced without bringing it into a theory from the outside, if $h$ is grasped as a new fundamental physical constant. This problem will be discussed later in Sec. 4 of the present chapter, where the Bohr model of atoms is treated.

§3. The quantum of action and periodic and aperiodic molecular processes

There were trials to investigate the physical implication of the quantum of action from the stand point of grasping it as a new fundamental physical constant, instead of attempts to deduce the quantum of action from the former fundamental physical constants. Planck presented "the hypothesis of the elementary quantum of action", in which the phase space of the resonator is supposed to consist of the elementary domains of area $h$, to explain the quantum nature of the energy of the resonator. Sommerfeld attempted to classify periodic phenomena such as X-ray production and photoelectric effect, with the condition imposed on the action integral, $\int_0^\tau Ldt = h/2\pi$ ($\tau$ being the time duration of the action).

Planck points out, in his lecture "The Laws of the Thermal Radiation and the Hypothesis of the Elementary Quantum of Action"$^{11}$ given at the First Solvay Conference (autumn, 1911), that Liouville's theorem in classical dynamics leads to Rayleigh–Jeans' distribution formula. Though it is well

$^{10}$Cf. p. 60, Vol. I.

known that the equipartition law of energy is deduced from this theorem, here Planck emphasizes the following point. That is, for an extremely short wave length $\lambda$, while in Planck's formula the energy density $\rho_\lambda$ of the thermal radiation becomes infinitely small with an exponential factor, in Rayleigh–Jeans' formula it becomes infinitely large increasing as $\lambda^{-4}$. Planck argues that "this is based on the fact that with vanishing wave length, that is, with increasing frequency, the number of the independent variables or of the degrees of freedom that corresponds to a specified spectral interval $d\lambda$ increases without limit".

Planck therefore thinks that certain restrictions should be introduced about the permissible values of the Lagrange variables $q$ and $p$, so that these quantities are expected to change stepwise only, or to be coupled to each other to some extent. By such restrictions, the number of the mutually independent elementary domains of identical probability is to be decreased, and this aim can be achieved by assuming the size of the elementary domain not to be infinitely small, but to be limited. Thus, Planck assumes

$$\int \int dp dq = h. \quad (10)$$

This is "the hypothesis of the elementary quantum of action".

Planck takes the case of a linear electric oscillator as the simplest emitting body, and puts its energy $E$ as$^{12}$

$$E = \frac{1}{2} K q^2 + \frac{1}{2} L \left( \frac{dq}{dt} \right)^2, \quad (11)$$

where $q$ is the electric dipole moment of the oscillator. The frequency is given by

$$\nu = \frac{1}{2\pi} \sqrt{\frac{K}{L}}.$$

With the use of $p = Ldq/dt$, $E$ is rewritten as

$$E = \frac{1}{2} K q^2 + \frac{1}{2L} p^2. \quad (12)$$

For the elementary domain of identical probability, Planck puts

$$h = \int \int_{E+\varepsilon} dp dq. \quad (13)$$

$^{12}$Cf. Eq. (4), Sec. 1, §1 of the last chapter.
This integral represents the area between two ellipses $E = \text{const.}$ and $E + \varepsilon = \text{const.}$ in the $q-p$ plane. Therefore, Eq. (13) becomes

$$h = 2\pi \sqrt{\frac{L}{K}} \varepsilon = \frac{\varepsilon}{\nu},$$

from which it turns out that there exists the elementary energy quantum

$$\varepsilon = h\nu.$$  \hspace{1cm} (15)

Planck emphasizes, thus, that in his opinion "the quantum hypothesis is not a hypothesis on the energy, but a hypothesis on the action".

On the other hand, Sommerfeld emphasizes that it is in aperiodic elementary processes that the quantum of action plays an important role, in a 1911 paper "On the Structure of $\gamma$ rays"$^{13}$ and in lecture at the First Solvay Conference "The Meaning of the Quantum of Action for Aperiodic Molecular Processes in Physics".$^{14}$ His viewpoint is characteristically stated in his following words — "One could build a mechanics that deals with only periodic circular motions. It would suffice for roughest needs of astronomy and for many questions of the machine technics. However it would give us only a very incomplete image about the laws of the mechanics".

It seems to him that "the theory of the energy quantum stays on a similar standpoint". As long as it is concerned only with periodic processes, it can meet the needs of the radiation theory and the main problems of specific heats. "About a more individual problem of single physical event it must make a stop", and therefore "one would have to search for a general standpoint". Considering so, Sommerfeld makes the assumption that "in every pure molecular process a determinate universal quantity of action is taken or given by the atom, namely the quantity

$$\int_{0}^{\tau} Ldt = \frac{h}{2\pi},$$

where $\tau$ is the duration of the action process". This assumption forms a basis for giving birth to the quantum condition, as will be seen in the next chapter. With the use of this assumption, Sommerfeld discusses the problems of X-ray


production by cathode rays, photoelectric effect, ionization potential, and γ-ray production. In the above equation, \( L \) is the "kinematic potential" given by

\[
L = T - V ,
\]

in terms of the kinetic energy \( T \) and potential energy \( V \).

(i) **X-ray production by cathode rays** Sommerfeld adopt the electromagnetic pulse theory due to Schuster–Stokes and Wiechert in that the produced X-rays are supposed to be the electromagnetic pulses accompanying the deceleration of a cathode ray particle (electron). The electromagnetic pulse theory of X-ray production explains only the polarized part of the produced X-rays. Over this part the unpolarized second part is superposed, which is considered to be the proper radiation or fluorescence of the target matter. The second part is most important for metals. The energy of X-rays of the polarized part \( E_{\text{pol}} \) is given by

\[
E_{\text{pol}} = \frac{e^2 u}{6\pi c^2} \frac{\beta}{\sqrt{1 - \beta^2}} ,
\]

where \( e \), \( u \) and \( \dot{u} \) are the electric charge, initial velocity and constant deceleration of the cathode ray particle, respectively, and \( \beta = \mu/c \). Sommerfeld then applies Eq. (16) to calculate the energy of the cathode ray particle \( E_k \), and compares \( E_{\text{pol}} / E_k \) with the experimental data.

He assumes that the potential energy of the electron due to the atom, which brakes the electron, is small compared with its kinetic energy, and puts

\[
V = 0 , \quad L = T .
\]

On the assumption of constant deceleration, he has

\[
\int_0^T T \, dt = \frac{m}{2u} \int_0^u u'^2 \, du' = \frac{1}{3} \frac{m}{2} u^2 \frac{\dot{u}}{\dot{u}} = \frac{\hbar}{2\pi} ,
\]

\( m \) being the mass of the particle. With the use of the relations \( \tau = u/\dot{u} \) and \( E_k = m u^2 / 2 \), this becomes

\[
E_k \tau = \frac{3}{2\pi} \frac{\hbar}{\pi} .
\]

\[\text{15}^\text{)} \] Nowadays, in usual text books on mechanics, \( L \) is called the Langrange function or Lagrangian.

\[\text{16}^\text{)} \] A. Sommerfeld, Ref. 13) of the present section.
Then, from Eqs. (18) and (20), he gets

$$\frac{E_{\text{pol}}}{E_k} = \frac{e^2}{9hc} \frac{\beta^2}{\sqrt{1 - \beta^2}}.$$  \hfill (21)

For $\beta = 0.4$, this gives

$$\frac{E_{\text{pol}}}{E_k} = 2.7 \times 10^{-4},$$  \hfill (22)

which agrees fairly with the experimental value $1.7 \times 10^{-4}$.

This experimental value is calculated from the value $E_{\text{pol}}/E_r = 1/4$ ($E_r$ being the total energy of the produced X-rays), which is obtained from an extrapolation of the X-ray intensities measured by Friedrich for glass at the azimuthal angles of $70^\circ$, $103^\circ$ and $133^\circ$, and $E_r/E_k = (2/3) \times 10^{-3}$ observed by Carter. In these two experiments $\beta$ is nearly 0.4, corresponding to an applied voltage of 40,000 V.

Agreement with the experimental value is improved if Eq. (20) is replaced by $E_k - \tau = h$, or $h/4$ is substituted for $h/2\pi$, in the right-hand side of Eq. (16), as is done in his previous paper.\(^\text{17}\) Further, if the energy, instead of the velocity, of the electron is assumed to decrease at a constant rate, it results that $E_k \tau = h/\pi$. The potential energy of the electron, which can be considered to arise from the atomic attractive force, probably corresponds to several volts, and may be neglected relative to the kinetic energy of about 40,000 volts. Sommerfeld makes these discussions about the result given in Eq. (22).

(ii) **Photoelectric effect** Sommerfeld calculates the action integral for the case of an electron under the influence of the electric field of a light wave, and shows that the kinetic energy of the photoelectron becomes $h\nu$ when the eigenfrequency $\nu_0$ of the atomic electron is equal to the frequency of the light.

If the atomic electron is assumed to be bound to the equilibrium point ($x = 0$) by an elastic force $-fx$, and is supposed to be forced by the incident wave $F = E \cos nt$, one has

$$m\ddot{x} + fx = eF = eE \cos nt; \quad n = 2\pi \nu,$$  \hfill (23)

\(^{17}\)A. Sommerfeld. Ref. 13) of the present section.
The action integral is given by

\[ W = \int_0^\tau \left( \frac{m}{2} \ddot{x}^2 - \frac{f}{2} x^2 \right) dt = \frac{m}{2} \ddot{x} - \frac{1}{2} \int_0^\tau x(m\ddot{x} + fx) dt. \]

On account of the equation of motion, Eq. (23), this can be rewritten as

\[ W = \frac{m}{2} \ddot{x} - \frac{e}{2} \int_0^\tau xF dt. \quad (24) \]

\( W \) is an oscillating function whose amplitude varies slowly with time. \( W \) gets the value \( \hbar/2\pi \) for the first time when \( W \) comes near its maximum value, because if it is not so \( W \) should get the value in the foregoing period. One therefore has \( dW/dt = 0 \) at the accumulation time \( t = \tau \), from which one deduces the relation \( T = V \), or \( \ddot{x}^2 = (f/m)x^2 \). Putting the eigenfrequency of the electron as \( \nu_0 = \sqrt{f/m}/2\pi \), one gets from Eq. (24)

\[ \frac{h}{2\pi} = \frac{\overline{T}}{2\pi \nu_0} - \frac{e}{2} \int_0^\tau xF dt, \]

that is, with \( n_0 = 2\pi \nu_0 \),

\[ T = h\nu_0 + \frac{1}{2} \nu_0 n_0 \int_0^\tau xF dt. \quad (25) \]

In the case of \( \nu = \nu_0 \), the “time virial” given by the integral in this equation can be shown to be negligible, as follows. In this case, one has for the solution to the equation of motion,

\[ x = \frac{eE}{2mn} t \sin nt, \]

in consideration of \( x = \dot{x} = 0 \) at \( t = 0 \) and \( n = n_0 \). Making use of this equation together with \( n\tau = \pm \pi/4 \), which comes from the condition that \( T = V \) at \( t = \tau \), one gets for the time virial

\[ \frac{en_0}{2} \int_0^\tau xF dt = \frac{(eE)^2}{32mn^2}. \]

This is \((n\tau)^{-2}/2\) times the value of \( T \) at \( t = \tau \), and is negligible because the number \( \nu\tau \) of the oscillations in the time \( \tau \) is great. One thus arrives at “Einstein’s law”

\[ T = h\nu_0 = h\nu. \quad (26) \]
The energy of the photoelectron does not depend on the intensity of the incident wave.

In the case of an incomplete resonance \( n \neq n_0 \), one calculates also the time virial with the use of the solution to the equation of motion

\[
x = -\frac{eE \cos nt - \cos n_0 t}{m \frac{n^2 - n_0^2}{2}}.
\]

and gets, in an approximation valid for \( n \) slightly different from \( n_0 \),

\[
T = h\nu_0 - \frac{(neE)^2}{2m(n^2 - n_0^2)^2} (\varepsilon - \sin \varepsilon); \varepsilon = (n - n_0)\tau.
\]

Now one has, besides \( T = V \) at \( t = \tau \), \( T = \bar{V} \) for the average of \( V \) in the last period of the oscillation of the electron. Calculating \( T \) from \( \bar{V} = (m/2)n_0\bar{x}_0^2 \), one obtains

\[
T = \frac{(n_0eE)^2}{2m(n^2 - n_0^2)} (1 - \cos \varepsilon).
\]

The comparison of Eq. (27) with Eq. (28) determines \( \varepsilon \) which gives \( \tau \).

This result shows that the emission of photoelectron is possible for \( n > n_0 \), and there is not such possibility generally for \( n < n_0 \). This corresponds to Stokes' rule for the phenomena of fluorescence. Photoelectrons may be emitted near \( n = n_0 - \pi/2\tau \), but deviation from Stokes' rule would be very slight because \( (n - n_0)/n \) is small on account of large \( \nu\tau \).

Sommerfeld calls the relation \( T = h\nu \) Einstein's law. About the fact that Eq. (27) is somewhat different from this relation, Sommerfeld argues that his result may be better than Einstein's law, because the experimental measurement of the electrostatic potential of Al irradiated for a long time with ultraviolet rays of high intensity shows that the photoelectric effect is clearly selective and influenced by the atomic eigenfrequency.

According to him, the numerical coefficient \( 1/2\pi \) in the equation \( \int_0^\tau Ldt = h/2\pi \) is chosen so as to get \( T = h\nu \) for \( n = n_0 \). If the numerical coefficient is chosen so as to yield \( T = h\nu \) for the smallest \( \tau \), the equation turns into \( \int_0^\tau Ldt = h/4 \). Sommerfeld "has not yet been able to hit on" a conclusive decision of these alternatives, as he states.
From the condition $T = h\nu$ at $t = \tau$, one has for the amplitude of the electron at that time

$$x = \sqrt{\frac{h}{mn\pi}}.$$ (29)

Sommerfeld says that this gives a new relation of the quantum of action to molecular quantities, and is similar to what was pointed out by Hass. Sommerfeld stresses, however, “I would not go further to see the true origin of $h$ in this connection”, and “I would on the contrary prefer the opposite standpoint not to explain $h$ from the molecular dimension, but to look at the existence of the molecule as a function and result of the existence of the elementary quantum of action”.

In relation to our discussions given above in §2 on Haas’s work, it will be made clear in later sections that this thinking of Sommerfeld is a foresighted one. Sommerfeld stresses, also, that “an electromagnetic or mechanical ‘explanation’ of $h$ seems just as little advisable and promising as a mechanical ‘explanation’ of the Maxwell equations”, and “it will be much effective to pursue the $h$-hypothesis in its various consequences and thereby to trace back other phenomena”.

(ii) Ionization potential difference Sommerfeld considers the case of the ionization of a molecule by an incident electron, and calculates the ionization potential difference by applying the hypothesis of the quantum of action to the electron. The ionization process is similar to the photoelectric effect, but is induced differently from it by an aperiodic electric field, instead of a periodic one. This process could be calculated in a way similar to that for the photoelectric effect, but here the incident electron is directly treated to show only the outline of the theory.

Supposing that the incident electron with initial velocity $u$ moves with a constant deceleration over a distance $l$ to stop, one has from Eq. (20)

$$\frac{1}{3} T\tau = \frac{h}{2\pi},$$ (30)

on the assumption that $V = 0$. In this equation, $T$ is the initial kinetic energy of the electron, and $\tau$ is given by

$$\tau = \frac{2l}{u}.$$ (31)
\( \tau \) takes the maximum value

\[
\tau = \frac{4\sigma}{u}
\] (32)

when \( l = 2\sigma \), \( \sigma \) being the range of the molecular force. The minimum energy required for the ionization process is therefore given by

\[
m_u\sigma = \frac{3}{4\pi}h.
\] (33)

If \( \sigma \) is taken to be \( 10^{-8} \) cm (He), one has \( u = 1.8 \times 10^{-8} \) cm/s, and thus the minimum energy required for the ionization is \( T = 1.5 \times 10^{-11} \) erg. This corresponds to the electric potential difference

\[
U = \frac{T}{e} = 9.3 \text{ V}.
\]

The values of \( U \) given by Townsend\(^{18}\) are 14.6 V for He, 17.3 V for A, 26 V for \( A_2 \), etc.

On the assumption that \( V = \text{const.} \) in the time \( \tau \), one has, instead of Eq. (30),

\[
\left( \frac{1}{3}T - V \right) \tau = \frac{h}{2\pi},
\]

which gives, instead of Eq. (33),

\[
m_u\sigma \left( 1 - \frac{3V}{T} \right) = \frac{3h}{4\pi}. \] (34)

If Townsend’s values are used for \( T \) and \( \sigma \), \( V \) takes the values 1.0 eV for He, 3.3 eV for A and 4.5 eV for \( A_2 \), etc.

Sommerfeld states that it is difficult to say what will happen when the electron is incident with a kinetic energy below the minimum value deduced above, because the ionization can not occur to stop the electron. “The electron will be reflected or go unimpededly through the molecule? Both do not seem probable”.\(^{19}\) He says also that such a rough use of the molecular size as is taken above is “not in accordance with our standpoint that the knowledge of


\(^{19}\) The pointing-out of this contradiction is important. This becomes solved later when the wave nature of the electron is discovered.
molecular properties should be substituted by the hypothesis of the quantum of action”.

In the case of the photoelectric effect, Sommerfeld used the model that the electron in the molecule is attracted to its equilibrium point by an elastic force. In the case of the ionization potential difference, however, he did not use any molecular model but took into consideration the incident electron. In the case of X-ray production by cathode rays, some atomic model would also be required to treat the unpolarized part of the radiation which is proper to the atom. A substantialistic knowledge of the atom was not enough to make the best use of Sommerfeld’s intention.

About γ-ray production, Sommerfeld adopted also the pulse theory as in the case of the X-ray production, and calculated the ratio of the energy of the γ-ray $E_\gamma$ to that of the incident electron $E_\beta$. However, the experimental data to be compared were not sufficient and he only got results more or less like the one that the magnitude of $\hbar$ required was of the order of $10^{-27}$ erg.s.

Both of Planck’s $\int \int dpdq = \hbar$ and Sommerfeld’s $\int Ldt = \hbar/2\pi$ developed into the quantum conditions the proposal of the Bohr model of atom. Descriptions of this process will be given later in Chap. 3. Sommerfeld’s view of $\hbar$ was adopted by Nicholson, who introduced the concept of the quantum nature of angular momentum. This will be discussed below in Sec. 3, §2 of the present chapter.

2. The Combination Rule of Spectral Lines

§1. Ritz’s theory of spectral lines

Ritz proposed in 1908 a “new combination rule” of spectral lines. This rule was a generalization of that conjectured by Rydberg before, and was confirmed by Ritz through his extensive analyses of the experimental results to predict unknown series of spectral lines. Rydberg supposed in 1890 that the wave number $\nu$ of a spectral line might be given by the formula

$$\nu = N \left\{ \frac{1}{(m+s)^2} - \frac{1}{(n+a)^2} \right\}; \quad (m,n = \text{integer}).$$

Ritz endeavoured to understand the combination rule of Rydberg in the framework of the classical theory. He devised in 1903 a special mechanical model, and in 1908 invented his “atomic magnetic field” model.

---

In his 1903 paper "On the Theory of Series Spectra", Ritz considers the oscillation of a square plate of side length $a$. This is because Rydberg's formula contains two integers. Let the small displacement of the plate in its normal direction be $w$ and the mass per unit area of the plate be $\rho$. The kinetic energy of the plate is given by

$$\int \int \frac{1}{2} \rho \left( \frac{\partial w}{\partial t} \right)^2 \, do,$$

where $do$ is the infinitesimal area element of the plate. Ritz expresses the normal displacement of the plate making use of Dupin's indicatrix. Let $x$, $y$ be the coordinates with origin at the center of the plate and along directions parallel to the two mutually rectangular sides of the plate, and $\xi$, $\eta$ be the diagonal coordinates. $x$, $y$ are related to $\xi$, $\eta$ by $x = \xi + \eta$ and $y = \xi - \eta$. Defining

$$\nabla w = \frac{\partial^2 w}{\partial \xi \partial \eta} = \frac{\partial^2 w}{\partial x^2} - \frac{\partial^2 w}{\partial y^2},$$

Ritz makes a special assumption. That is, in terms of

$$X = |x - x'| - |x - x'|^2/2a,$$

$$Y = |y - y'| - |y - y'|^2/2a,$$

$$\Phi = u \int \int XY \nabla' wdo',$$

together with a constant $u$, he assumes that $\Phi^2$ gives the density of the potential energy, ($\nabla' w$ denoting $\nabla w$ with respect to $x'$ and $y'$).

From the Hamilton principle

$$\int \int \left[ \frac{1}{2} \rho \left( \frac{\partial w}{\partial t} \right)^2 - \Phi^2 \right] \, dodt = \text{Min},$$

---

21) W. Ritz, Ann. d. Phys. 12 (1903), 264

22) For a curved surface given by $w = f(\xi, \eta)$, Dupin's indicatrix takes the form

$$(\partial^2 f / \partial \xi^2) \xi^2 + 2(\partial^2 f / \partial \xi \partial \eta) \xi \eta + (\partial^2 f / \partial \eta^2) \eta^2 = \epsilon,$$

with an infinitesimal parameter $\epsilon$. This gives the intersection between the curved surface and a plane parallel to and near a contact plane of the curved surface.
he derives the fundamental equation

$$\rho \frac{\partial^2 w}{\partial t^2} + 2u \nabla \varphi = 0,$$

(37)

with \( \varphi = \int \Phi XY \, d\alpha \) and the boundary conditions

$$\frac{\partial w}{\partial n} = 0, \int_{-a}^{+a} w \, dx = 0, \int_{-a}^{+a} w \, dy = 0, \frac{\partial \varphi}{\partial n} = 0,$$

(38)

where \( n \) denotes a normal to a side of the square. The form of \( \Phi \) indicates that the displacement at the point \((x', y')\) exerts an "action at a distance" on the point \((x, y)\), which is proportional to the degree of the deformation \( \nabla' w \) and the quantity \( XY \).

The fundamental equation is rewritten as

$$\rho \frac{\partial^{10} w}{\partial t^2 \partial x^4 \partial y^4} + 32u^2 \nabla \nabla w = 0,$$

(39)

whose solutions satisfying the boundary conditions and corresponding to eigen-oscillations are given by

$$w = A \sin(\nu t + c) \cos \left( \frac{m \pi x}{a} \right) \cos \left( \frac{n \pi y}{a} \right),$$

(40)

in terms of arbitrary constants \( A \) and \( c \), two integers \( m \) and \( n \) and \( \nu \) defined by\(^\text{23)}\)

$$\rho \nu^2 - 2u^2 \left( \frac{1}{m^2} - \frac{1}{n^2} \right)^2 \frac{16a^4}{\pi^4} = 0.$$  

Putting

$$N = \sqrt{\frac{32 \rho}{u}} \frac{a^2}{\pi^2},$$

\( \nu \) is given by

$$\pm \nu = N \left[ \frac{1}{m^2} - \frac{1}{n^2} \right]; \quad m, n = \pm 1, \pm 2, \pm 3, \ldots.$$  

(41)

Ritz considers that these correspond to the Balmer series.\(^\text{24)}\)

\(^{23)}\) \( \nu \) used here by Ritz denotes the angular frequency.

\(^{24)}\) Cf. Sec. 1, § 2, p. 63, Chap. 2, Vol. I. Experimental confirmation of the series with \( m = 1 \) (the Lyman series), \( m = 3 \) (the Ritz–Pashen series), \( m = 4 \) (the Brackett series), and \( m = 5 \) (the Pfund series) took place respectively in 1906, 1908, 1922, and 1924.
If $\Phi$ is defined in another way by the partial differential equation
\[
\frac{\partial^4 \Phi}{\partial x^2 \partial y^2} = 4\nu w, \tag{42}
\]
the same equation as Eq. (39) is obtained from
\[
\iint \left[ \frac{1}{2} \rho \left( \frac{\partial w}{\partial t} \right)^2 - (\nabla \Phi)^2 \right] \, dt = \text{Min}. \tag{43}
\]
Taking for example
\[
\Phi = u \iint w |x - x'| Y \, dx' \, dy'. \tag{44}
\]
Because the boundary condition for $\Phi$ can be chosen arbitrarily, the solutions take the form
\[
w = A \sin(\nu t + c) \sin \left( \frac{(m + 1/2)\pi x}{a} \right) \cos \frac{n\pi y}{a}, \tag{45}
\]
\[
\pm \nu = \sqrt{\frac{32 \, u a^2}{\rho \, \pi^2}} \left[ \frac{1}{(m + 1/2)^2} - \frac{1}{n^2} \right];
\]
\[
m = 0, \pm 1, \pm 2, \ldots; \quad n = \pm 1, \pm 2, \ldots. \tag{46}
\]
According to Ritz, these correspond to the principal series and the sharp series \textit{(die zweite Neben serie)} of hydrogen atom.\(^{25}\)

For the series of other elements, it is difficult to make calculations similar to those done above, because nothing is certain other than the fact that some constant difference in the wave number is recognized in the sharp and diffuse series. Ritz states, however, that we are led to the assumption that the spectral series of all the elements come from the oscillating system treated above, with boundary conditions varying from one series to another. He supposes, thus, that the forms of Eqs. (39) and (42) are changed also for other elements.

$\nu$ is then assumed to take the following form:
\[
\nu^2 = \left[ \frac{1}{k^2} - \frac{1}{l^2} \right]^2 \frac{32}{\rho} u^2, \tag{47}
\]
\(^{25}\)For the German and English names of the spectral series, confer sec. 1, \S 2, p. 64, Chap. 2, Vol. I.
under a quite general kind of boundary condition. In this equation, \( k \) or \( l \) is a solution of a transcendental equation of the form \( \tan ka = \varphi(k) \) or \( \tan la = \varphi(l) \), respectively. The iterative solution of the transcendental equation, with the development of \( \varphi(k) \) into the power series of the inverse of \( k \), gives to \( ka \) a formula of the type

\[
ka = m\pi + c_0 + \frac{c_1a}{m\pi + c_0} + \frac{c_2a^2}{(m\pi + c_0)^2} + \cdots,
\]

where \( m \) is an integer. Making use of this and a similar one for \( la \) with another integer \( n \) and neglecting the fractional terms in them, Ritz obtains for \( \nu \) the formula

\[
\pm \nu = N \left[ \frac{1}{(m + \mu)^2} - \frac{1}{(n + \nu)^2} \right]; \quad N = \sqrt{\frac{32}{\rho} \mu a^2 / \pi^2}. \tag{47}
\]

He considers that this is in accordance with the Rydberg formula.

Using the formula for \( ka \) given above and taking a fixed value for the integer \( n \), he obtains the following formula for a spectral series:

\[
\pm \nu = A - \frac{N}{[m + \mu + b/(m + \mu) + \cdots]^2}. \tag{48}
\]

The constant \( N \) in this formula is universal for all the spectral series. Ritz considers that this formula is better expressed in the form

\[
\pm \nu = A - \frac{N}{[m + \mu + \beta(A - \nu)]^2} \tag{49}
\]

for the purpose of improving the degree of approximation.

On the basis of these formulae, Ritz analyzes the spectral lines of \( \text{H, He}_1, \text{He}_2, \text{Li}, \text{Na}, \text{Ka}, \text{Rb}, \text{Cs}, \text{Ag}, \text{O}_1, \text{Mg}, \text{Ca}, \text{Sr}, \text{Tl}, \text{Zn}, \text{Cd} \) and \( \text{Hg} \), and thereby determines the values of the constants in the formulae.

It has already been mentioned in Vol. I that the greatest difficulty in dynamical models of spectral lines lies, as was pointed out by Rayleigh,\(^{26}\) in the fact that any dynamical model deduces a formula for the square, \( \nu^2 \), of the frequency \( \nu \), while the observational results of spectral lines determine the frequency \( \nu \) itself. For this reason, Ritz arrives only at the result that \( \nu^2 \) is expressed as the square of the difference between two terms, in spite of his

\(^{26}\)Lord Rayleigh, Phil. Mag. 44 (1897), 356; (cf. p. 73, Vol. I).
“complicated assumptions of which physical meanings could not be found”,
according to his words in a 1908 paper to be discussed below.\(^{27}\)

§2. Ritz’s “atomic magnetic field” model

Ritz considers, in his 1908 paper “Atomic Magnetic Field and Series Spectra”\(^{28}\) that every atom has a magnetic field due to a series of bar magnets, and discusses the motion of an electron in the magnetic field. He says, referring to Weiss’s paper on the theory of molecular magnetic field,\(^{29}\) that “it is almost inevitable to assume the existence of strong atomic magnetic fields”, and attempts to explain spectral series and the anomalous Zeeman effect in terms of pure magnetic forces.

If there is a charged particle with the electric charge \(e\) and mass \(m\) in a homogeneous magnetic field \(H\) parallel to the \(z\) axis, it holds that

\[
\frac{m}{c} \frac{d^2x}{dt^2} = \frac{eH}{c} \frac{dy}{dt}, \quad \frac{m}{c} \frac{d^2y}{dt^2} = -\frac{eH}{c} \frac{dx}{dt}, \quad \frac{m}{c} \frac{d^2z}{dt^2} = 0;
\]

\[
x = A \sin \nu(t - t_0), \quad y = A \cos \nu(t - t_0), \quad z = z_0 + Bt,
\]

where \(A, B, t_0\) and \(z_0\) are arbitrarily constants and \(\nu = eH/mc\). If the magnetic field \(H\) is not homogeneous, it suffices to consider the case of small \(x, y\) and \(z - z_0\), or a case in which the particle is attracted toward the origin by a comparatively weak elastic force.

The magnetic field \(H\) is assumed to be generated by a series of \(n\) elementary magnets in a line, each of length \(s\) and pole strength \(\mu\). If the particle is at a point on the prolonged line of the series of magnets, the magnetic field is given by

\[
H = \mu \left[ \frac{1}{r^2} - \frac{1}{(r + ns)^2} \right],
\]

where \(r\) is the distance between the particle and the nearest terminal of the series of magnets. \(\nu\) is therefore given by

\[
\nu = \frac{\mu e}{mc} \left[ \frac{1}{r^2} - \frac{1}{(r + ns)^2} \right] \quad (50)
\]

\(^{27}\)In expressing Eq. (48) in the form of Eq. (49), only \(+\nu\) in \(\pm\nu\) is used in \((A - \nu)\) in the denominator.

\(^{28}\)W. Ritz, Ann. d. Phys. 25 (1908), 660.

\(^{29}\)P. Weiss, Compt. Rand. 143 (1906), 1136; 145 (1907), 9 Dec. et 30 Dec.
Ritz considers that this resembles closely Balmer’s formula, Rydberg’s formula and Ritz’s formulae Eqs. (48) and (49), so that by putting \( r = as \) he rewrites it as

\[
\nu = \frac{\mu e}{s^2mc} \left[ \frac{1}{a^2} - \frac{1}{(n + a)^2} \right] ; \quad n = 1, 2, 3, \ldots .
\]  

(51)

If the length of the first elementary magnet in the series is assumed to be \( s(1 + \varepsilon) \), Eq. (51) takes the form of Rydberg’s spectral series formula

\[
\nu = N \left[ \frac{1}{a^2} - \frac{1}{(n + b)^2} \right] ; \quad b = a + \varepsilon .
\]

Fixing \( a \) and varying \( \varepsilon \), one gets spectral series other than Rydberg’s, which have the same limiting frequency as that of Rydberg’s series. Ritz considers that these series correspond to the well known sharp and diffuse series and doublet and triplet terms of the principal series. According to him, from Eq. (49), which expresses the experimental values about the elements other than hydrogen, it is seen that the distances between the two magnetic poles and the charged particle do not obey the simple additive law shown above.

Ritz mentions that the magnetic field due to a rotator with surface charges is equivalent to the magnetic field due to a bar magnet with poles along the axis of the rotator, and discusses the possibility that the atomic magnetic field is caused by such a charged rotator. If the rotator is a cylinder, the magnetic poles come very near to the terminal sections of the cylinder, so that the density of the surface charges becomes very large there. Near the poles the charged particle makes a rotating motion due to the magnetic field. A line of \( n \) rotators of a common axis and form, and of alternative signs of the surface charges and directions of rotation, makes therefore a system that is very stable and almost equivalent to a bar magnet of the same length as the \( n \) rotators.

It may be said to be interesting that Ritz was able anyway to derive a formula similar to the law of spectral lines. However, for Ritz’s result to coincide with the law of spectral lines, the length of the bar magnet or the number of the elementary magnets should be different for each spectral line of a series. It should be assumed that there are magnets consisting of various numbers of elementary magnets. In the case of the model of the oscillating plate, he attributed the variety of spectral lines to the variety of boundary conditions for the function \( \Phi \) obeying a partial differential equation. In the case of the atomic magnetic field model, the variety of spectral lines is attributed to
the variety of lengths of the bar magnet. No relation is investigated between
the variety of the bar magnets and the atomic structure.

Ritz was more interested in the deduction of the frequency formulas than
the search for substantialistic clues to the atomic structure. Ritz's models
should thus be said to be analog models. The model of the oscillating plate
cannot be related to the emission of light because it remains a mechanical
model, though it could give frequencies obeying a formula similar to that for
spectral lines. The atomic magnetic field model is a kind of electromagnetic
model and has a substantialistic character in the sense that it might be related
to light emission. It has the stability of atomic electrons, but not, as it is, the
electric neutrality of the atom.

In any way, Ritz's models possessed the difficulties mentioned above and
could not become models of concrete existence. Substantialistic attempts such
as were made in these models did not, however, come to nothing. The model
of oscillating plate discussed in §1 has an aspect of the oscillation of the ring
of electrons of Nagaoka–Nicholson type, and the atomic magnetic field model
has an aspect of describing the classical appearance of the modes of oscilla-
tion of atomic electrons. With these aspects as the moment Ritz abstracted
his models, and pushed forward his phenomenological analyses of complicated
spectral lines. The models played the role of giving the clue to this procedure.
His systematization of a wide range of observational results with the use of his
spectral formula was effective in bringing out the combination rule that will
be discussed below in §3.

Now, in his discussions on the Zeeman effect, Ritz calculates the motion of
a charged particle in an external magnetic field $H$ in addition to the atomic
magnetic field $H_0$, making use of the Lagrange equations of motion. In this
calculation, the coordinates of the charged particle in the plane normal to $H_0$
are chosen as the Lagrange coordinates, on the assumption that the particle
makes an oscillation of small amplitude in the plane normal to $H_0$ while $H_0$
rotates about $H$. There appear in the kinetic energy, the zenithal and az-
imuthal angles $\theta$ and $\varphi$, respectively, of $H_0$ in the fixed system of coordinates
having the z-axis in the direction of $H$, together with their time derivatives.
The potential energy is not taken into account.

If $H_0$ rotates normally to $H$ with a constant angular frequency $\omega$, the
angular frequency $\nu$ of the particle is given by

$$\nu = \nu_0, \quad \nu_0 \pm \omega; \quad (\nu_0 = eH/mc).$$
Ritz supposes that this corresponds to the case of Lorentz–Larmor. In the Lorentz–Larmor theory of electron, the angular frequency $\nu_0$ of the electron making an eigenrotation splits into $\nu_0, \nu_0 \pm eH/2mc$ upon the action of an external magnetic field $H$. Ritz says that “the interpretation of the Zeeman effect is essentially changed” by the assumption of the atomic magnetic field.

The reason why he does so is to explain the anomalous Zeeman effects, in which quintet, nonet, quartet and octet terms are seen. Ritz assumes that the zenithal angle $\theta$ rotates with an angular frequency $\omega$ and the azimuthal angle $\varphi$ with an angular frequency $\omega'$. Noting the fact that $\omega'/\omega$ should take a rational value for the motion to be periodical, he makes calculation, using Fourier series developments in $\cos \theta$ and $\varphi$, to get the quintet and nonet terms in the form $\nu_0 \pm m\omega (m = 0, 1, 2, \ldots)$. He considers also the case of $\omega' = eH/mc$, to obtain the quartet and octet terms in the forms $\nu_0 \pm \omega, \nu_0 \pm \omega'$ and $\nu_0 \pm \omega \pm \omega'$. In reality, the anomalous Zeeman effect becomes clearly understood after the discovery of the quantization of angular momentum and the discovery of the spin.

§3. Ritz’s combination rule

As is mentioned at the beginning of §1, Ritz proposed in 1908 a “new combination rule”. This was done in his paper “On a New Law of Spectral Series”. In this paper, Ritz abbreviates the formula (49) for series spectra given by him in 1903 as

$$ (m, \alpha, \beta) = \frac{N}{[m + \alpha + \beta(A - \nu)]^2}. $$

(52)

In accordance with Eq. (48) which results from the iteration approximation to $A - \nu$, this is rewritten as

$$ (m, \alpha, \beta) = \frac{N}{[m + \alpha + \beta N/m^2 - 2\alpha\beta N/m^3 + \cdots]^2}. $$

(53)

In these expressions, $N$ denotes the Rydberg constant 109675.0 cm$^{-1}$. $(m, \alpha, \beta)$ later comes to have the name of spectral term. The wave number of any spectral line is thus given by the difference between two spectral terms. $\alpha$ and $\beta$ are constants characterizing a spectral series.

---

30) Cf. Sec. 3, §1 and §3, Chap. 2, Vol. I.

31) W. Ritz, Phys. Zeit. 9 (1908), 521.
Ritz rewrites \( \alpha, \beta \) as \( \alpha, \delta \) for the diffuse series (\textit{die erste Neben serie}), and as \( s, \sigma \) for the sharp series (\textit{die zweite Neben serie}), and as \( p_1, \pi_1 \) and \( p_2, \pi_2 \) for the principal series \((p_1 > p_2, \pi_1 \approx \pi_2)\). In the diffuse series, there are satellite lines that lie near the respective main lines with a constant difference in wave number. To express this fact, he introduces another pair of constants \( d', \delta'(\delta \approx \delta') \).

Thus, in the new combination rule, these lines are given by the following formulae.

The principal series:

\[
\pm \nu = \begin{cases} 
(1.5, s, \sigma) - (m, p_1, \pi_1) & \text{smaller } \lambda \\
(1.5, s, \sigma) - (m, p_2, \pi_2) & \text{larger } \lambda 
\end{cases} \quad m = 2, 3, 4, \ldots
\]

The sharp series:

\[
\pm \nu = \begin{cases} 
(2, p_1, \pi_1) - (m, s, \sigma) & \text{larger } \lambda \\
(2, p_2, \pi_2) - (m, s, \sigma) & \text{smaller } \lambda 
\end{cases} \quad m = 1.5, 2.5, 3.5, \ldots
\]

The diffuse series:

\[
\pm \nu = \begin{cases} 
(2, p_1, \pi_1) - (m, d, \delta) & \text{main line, larger } \lambda \\
(2, p_2, \pi_2) - (m, d, \delta) & \text{satellite, } m = 3, 4, 5, \ldots \\
(2, p_2, \pi_2) - (m, d', \delta') & \text{main line, smaller } \lambda 
\end{cases}
\]

The constant difference between the doublet lines is given by \( \nu_1 = (2, p_1, \pi_1) - (2, p_2, \pi_2) \).

Ritz shows that “from known spectral series of an element new series can be deduced, by which almost all the series and lines, discovered in alkalis shortly before by Lenard, Konen and Hagenbach, Moll, Bergmann and others, are in particular well represented, without introducing any new constant”, and that “the new combination rule applies also to other spectra, particularly to He and the alkali earths”.

He points out that the following becomes clear from the previous observational results.

(1) By changing 1.5 to 2.5, 3.5, \ldots, one obtains new lines, many of which are observed, “so that one can write the above formulae with two ordinal numbers \( m, n \), as Rydberg already supposed”.
The Combination Rule of Spectral Lines

(2) For every \((m, \alpha, \beta)\) there exists a smallest \(m\), which corresponds to the ground line. By giving a still smaller value, one expects a still stronger line. In reality such a line is not observed. The line with \(m = 2\) in the diffuse series is not found as far as the observation covers.

(3) By making a new combination \(\nu = (1.5, s, \sigma) - (3, d, \delta)\), one gets the not so intensive lines which are observed for He, K and Rb and fall into the infrared region for Cs and alkali earths.

(4) The combinations \(\nu = (2, p_1, \pi_1) - (m, p_1, \pi_1)\), \(\nu = (2, p_2, \pi_2) - (m, p_2, \pi_2)\) are also detectable.

(5) By making the combination \(\nu = (3, d, \delta) - (m, p_1 - p_2, \pi_1 - \pi_2)\), \(m = 4, 5, 6, \ldots\), one obtains the infrared series found by Bergmann for K, Rb and Cs, and unknown corresponding lines for Li, Na and He. For Rb and Cs, because the diffuse series has satellite lines, one obtains the second combination \(\nu = (1.5, s, \sigma) - (3, d, \delta)\) too. This should be near the first with a constant separation, and is in fact observed for Cs, but is not separated for Rb. \(m = 4\) is the ground number of \((m, p_1 - p_2, \pi_1 - \pi_2)\) and \(\pi_1 - \pi_2 \approx 0\).

(6) There exists the combination \(\nu = (2, p_i, \pi_i) - (m, p_1 - p_2, \pi_1 - \pi_2)\), and

(7) \(\nu = (4, p_1 - p_2, \pi_1 - \pi_2) - (m, p_1 - p_2, \pi_1 - \pi_2)\). \(m\) can undoubtedly take large values, but the corresponding lines are generally too weak to detect.

For alkali earths, there are not extensive studies in the infrared region, and the principal series are not observed, so that tests of the combination rules in the sense mentioned above must be postponed. The following circumstances show, however, that here the combination rule also remains valid.

For the triplet series with two satellite lines in the diffuse series, one introduces the notations \((m, p_3, \pi_3)\) and \((m, d'', \delta'')\) with \(p_1 > p_2 > p_3\). Neglecting \(\pi_1 - \pi_2\) and \(\pi_2 - \pi_3\) in the first approximation, one then calculates \(p_1 - p_2\) and \(p_2 - p_3\) from the limits of the diffuse series, because the principal series are not observed. It turns out that \(p_1 - p_2\) is about twice \(p_2 - p_3\), so that the differences in wave number which Rydberg denoted by \(\nu_1\) and \(\nu_2\) are written as

\[\nu_1 = (2, p_2, \pi_2) - (2, p_1, \pi_1),\]
\[\nu_2 = (2, p_3, \pi_3) - (2, p_2, \pi_2).\]

As is known, the quantity \(\nu_1/\text{(atomic weight)}^2\) does not vary very much for the chemically alike elements, but does for the other elements under study.
However, the quantity \( (p_1 - p_2)/(\text{atomic weight})^2 \) does not vary very much for all the elements under study. The fluctuation of this quantity may arise from neglect of \( \pi_1 - \pi_2 \) which produces large errors for heavy elements.

Further, about the pairs of intensive ultraviolet lines of the alkali earths, setting up similar equations for the sharp series, one finds \( \nu_1 \) that is about twice \( \nu_1 \), and calculates from the limits of these series the difference \( p_1' - p_2' \) that is different by about 15\% from \( p_1 - p_2 \) of the triplet series discussed above. More exact calculations would give mutually identical values for these quantities, corresponding to the "combination rule of constants".

For Ca and Sr, there are side series (Nebenserien) of close triplets discovered shortly before. \( p_1' - p_2' \) and \( p_2' - p_3' \) obtained from the limits of the series are equal to \( d - d' \) and \( d' - d'' \), respectively, calculated from the main lines of the diffuse series, so that there exists a new combination.

These circumstances show that in many cases it is profitable to consider the constants of the series formulae, instead of the wave numbers.

In the alkali earths there are triplet lines besides the well known series. According to the combination rule, they are expected when \( (2, p_i, \pi_i) \) is connected with \( (m, \alpha, \beta) \). They belong to such series that terminate at the same places as the respective known series. But, they have only a few sufficiently intensive terms to be detected, and are often limited to their ground lines.

As the atomic weight increases the series becomes weaker, so that only lines corresponding to low ordinal numbers are observed. On the other hand, the number of different series increases (for example, the series of doublet and singlet lines besides that of triplet lines in the case of Mg). The number of the combinations increases too. One arrives finally at a spectrum, in which a large number of characteristic constant intervals in wave number are seen but no series is found.

After making the summary as described above, Ritz gave a detailed study of the spectra of the elements of \( H \), alkalis, He and alkali earths, to get the values of the constants \( s, p \) and others. \( (m, s, \sigma), (m, p, \pi) \) and \( (m, d, \delta) \) were later renamed \( ms-, mp- \) and \( md- \) terms respectively. About \( H \), Ritz considered that the line \( \lambda = 18751\text{Å} \) observed by Pascen would correspond to \( \nu = N(1/3^2 - 1/m^2) \), and supposed the existence of a series for which \( \nu \) is given by putting \( m = 4, 5, 6, \ldots \) in this formula. This series became afterwards confirmed by Pascen and named the Pascen series.

In this way Ritz arranged in detail, about the various spectral series of many elements, the combination rule of spectral lines conjectured by Rydberg.
This made it certain that spectral lines are given by combinations in difference of spectral terms.

The combination rule of the Rydberg–Ritz type is fundamentally different from the series of frequencies in the classical theory. The existence of spectral terms and their combination in difference has later become understood as the existence of stationary states and the energy-frequency relation in the Bohr model of the atom. This understanding has opened the way to further developments of the quantum theory.

This process has been so intricate that it has advanced in a spiral way, so as to say. The following part of the present volume is devoted to the analysis of this process.

3. The Quantum of Action and the Angular Momentum

§1. Nicholson's theory of spectra

As is mentioned at the beginning of the present chapter, when Bohr has proposed in his atomic model the existence of discrete stationary states due to the quantum nature of angular momentum, a good use has been made of the fact that Nicholson studied the spectrum of solar corona on the basis of an atomic model of the Nagaoka type with a nucleus, and pointed out the quantum nature of angular momentum in terms of the Planck constant.

In his 1911 paper "The Spectrum of Nebulium", Nicholson discusses this spectrum with the use of an atomic model which "closely conforms" to the Nagaoka model of the atom. Nicholson takes notice that there are nebular spectral lines which cannot be ascribed to hydrogen or helium, and some or perhaps all of them are supposed to come from an unknown atom, provisionally called nebulium, with the simplest possible type. According to him, this was communicated to British Association at Portsmouth. He says that "this paper is devoted to a preliminary examination of this hypothesis, and the results which are obtained appear to indicate that a dynamical system which can give rise to these spectral lines has been found".

"The main conception involved in the structure of these simple atoms is that of the nature of positive electricity. This is supposed to exist in small spherical volume distributions of uniform density, whose radius is small in comparison even with that of an electron, a reversal of the more generally accepted view",

and "the mass of these positive units is very large in comparison with that of an electron". In these words, Nicholson describes the constitution of a nebulium atom. It differs from the Nagaoka model of the atom in that it consists of fewer number of electrons. That is, "the atom of nebulium contains four electrons, each with a charge \(-e\) rotating uniformly at equal distances in a circle round a positive nucleus whose charge is \(4e\)."33) On this account, Nicholson says that "to Nagaoka's model of the atom these elementary simple substances closely conform".

The date of Nicholson's paper is "1911 October 25", but in his paper, Rutherford's paper which was published in the May 1911 issue of Phil. Mag. and in which the model of the atom with nucleus was confirmed, is not cited. For Nicholson's theory, it suffices to say that the origin of the principal concept of the atomic structure is not in the Thomson model but in the Nagaoka model.34)

Now, Nicholson first calculates the vibrational frequencies of the nebulium atom, and seeks their correspondence with the wavelengths of the main spectral lines. As a result he gets the value \(3.45 \times 10^{-8}\) cm for the radius of the nebulium atom, and then investigates the spectral lines of nebulium ions. He assumes that in a steady motion the electrons are in one plane, and considers mainly the vibrations perpendicular to the plane, for the reason that they appear to be the more important set for the visible spectrum. The major part of the necessary analysis is an application of that used in the Thomson model.

Let there be \(n\) electrons in a steady circular motion of radius \(a\) round a positive nucleus of charge \(\nu e\), and let the system be referred to cylindrical polar coordinates \((r, \theta, z)\), where \(z\) is perpendicular to the plane, with the origin at the center of the undisturbed orbit.

The force along increasing \(z\) exerted by the \(s\)th electron on the \(p\)th electron is given by

\[
(e^2/8a^3)(z_p - z_s)\csc^3 \psi; \quad \psi = |p - s|\pi/n,
\]

so that the force \(Z_p\) on the \(p\)th electron is

\[
Z_p = z_p D - \sum_{s=1}^{n-1} z_{p+s} D_s,
\]

33) Nicholson wrote "a positive nucleus". Though in some books it is said that Bohr used for the first time the word "nucleus" about the atom. Nicholson used it before Bohr.

34) On this point, confer p. 76, §2 of the present section.
where

\[ 8a^3 D = e^2 \sum_{s=1}^{n-1} \csc^2 s\pi/n; \quad 8a^3 D_s = e^2 \csc^2 s\pi/n. \]

Let \( \zeta \) be the displacement of the positive charge. The corresponding force on the electron due to the positive charge and the attraction of the electrons on the positive nucleus tending to increase \( \zeta \) are given respectively by

\[ -\nu e^2(z_p - \zeta)/a^3 \quad \text{and} \quad \sum_{s=1}^{n} \nu e^2(z_s - \zeta)/a^3. \]

If the motion has a period of \( 2\pi/q \), \( q \) being the angular frequency common to all quantities, there result from the equations of motion of the electrons and the nucleus

\[ -mq^2 z_p = Dz_p - \sum_{s=1}^{n-1} D_s z_{p+s} - (\nu e^2/a^3)(z_p - \zeta), \tag{53} \]

\[ \nu e^2 \zeta = \alpha \sum_{s=1}^{n} z_s a^3; \quad \alpha = \frac{(\nu e^2/a^3)^2}{(\nu e^2/a^3) - Mq^2}, \]

where \( m \) and \( M \) are the masses of the electron and the nucleus respectively. Elimination of \( \zeta \) leads to

\[ z_p \left\{ \left( \frac{\nu e^2}{a^3} \right) - Dmq^2 + \alpha \right\} + \sum_{s=1}^{n-1} (a + D_s) z_{p+s} = 0, \]

from which the period equation is deduced:

\[ mq^2 = \frac{e^2(8\nu + P_\kappa - P_0)}{8a^3}. \tag{54} \]

In this equation,

\[ P_\kappa = \sum_{s=1}^{n-1} \cos \left( \frac{2ks\pi}{n} \right) \csc^3 \left( \frac{s\pi}{n} \right), \quad \left( \frac{e^2}{8a^3} \right) P_0 = D. \]

\( \kappa \) is called the class of the vibration, and takes values \( \kappa = \pm(0, 1, 2, \ldots, n-1) \). The sign of \( \kappa \) is irrelevant to \( P_\kappa \) and \( q^2 \).
Nicholson says, here, on citing Schott's paper,\(^{35}\) that vibrations of class 3 or higher will not ordinarily appear in the visible spectrum. Nicholson considers that the mechanism of production of spectrum contemplated in his paper is not quite identical with that in Scott's paper treating ionized rings, but Schott's conclusions in this respect should nevertheless apply to the present case.

Schott's paper cited above is the one that we have already discussed in Sec. 6, §3, Chap. 2, Vol. I, and treats the intensities of lines produced by vibrations of a ring of electrons. As we have mentioned there, he showed that in the case of a single ring the spectral lines intensive enough to be detected by a photographic plate are only those of \(\kappa = 0, \pm 1\), while those of \(\kappa = \pm 2\) may be seen in some circumstances.

Nicholson attempts to solve this difficulty by considering the equilibrium among neutral atoms and positive as well as negative ions of nebulium. For the nebulium ion with two negative charges \((\nu = 4, n = 6)\), the vibration of \(\kappa = \pm 3\) has an imaginary frequency, so that this mode is unstable. However, because this mode has a very small amplitude, he supposes that "the atom with 6 electrons might persist for a considerable time, and be constantly formed again, if broken up, by the capture of stray electrons". Thus, Nicholson thinks that "it is quite possible that the lower modes of this system should give rise to spectrum lines in a mass of gas under high electrical excitment".

In his second paper on the solar corona, which will be taken up shortly below, Nicholson also thinks, about the instability of the system with \(\nu = n = 5\) for the vibration of \(\kappa = \pm 2\) in the plane, that "when an unstable vibration has resulted in the ejection of one or more electrons from a system, the new system can again take them up until the unstable modes arise again". He assumes therefore that "in a vibrating gas of this type, there would ultimately be a kind of equilibrium established in the interchange of electrons, and spectral lines might appear in correspondence with all positive or negative degrees of charge". The reason why Nicholson treats nebular and solar coronae consists in this point.

Now, returning to the problem of the neutral nebulium atom for which \(\nu = n = 4\), one has \(|\kappa| = 0, 1, 2, 3\). Because \(P_1 = P_3\) in this case, Eq. (54) becomes

\[
ma^3 q^2 = e^2 \left\{ 4, 4 - (2 + 4\sqrt{2})/8, 4 - \sqrt{2} \right\},
\]

\(^{35}\)G. A. Schott, Phil. Mag. \textbf{13} (1907), 190.
corresponding respectively to $|\kappa| = 0$, $|\kappa| = 1$ or 3, and $|\kappa| = 2$. Let $\omega$ be the angular velocity of the ring of electrons in steady motion. The balance between the centrifugal and electric centripetal forces exerted on the ring takes the following form in the general case of $n$ electrons in a ring of radius $a$

$$m\omega^2 = \left(\frac{e^2}{a^2}\right)\left(\nu - \frac{S_n}{4}\right); \quad S_n = \sum_{s=1}^{n-1} \csc\left(\frac{s\pi}{n}\right), \quad (56)$$

where $S_n$ expresses the effect of the repulsive forces exerted on one electron by the other electrons. Because $S_4 = 1 + 2\sqrt{2}$, one gets from Eq. (55) and Eq. (56) for $n = 4$

$$q/\omega = \{1.146533, 1, 0.849778\}. \quad (57)$$

In consideration of the rotational motion of the ring, Nicholson notes that "the frequency manifested to an observer is not $q$, but $q + \kappa\omega$". Denoting $|q + \kappa\omega|$ anew as $q$, one has for $-3 \leq \kappa \leq +3$

$$q/\omega = \{1.146533, 1.150222, 2, 2.849778, 4\}, \quad (58)$$

corresponding respectively to $\kappa = 0$, $\kappa = -2$, $\kappa = 1$ and $-3$, $\kappa = 2$, $\kappa = 3$. The term for $\kappa = -1$ is omitted because in this case one has $q\omega = 0$. Equation (58) expressed in the wave length $\lambda$, becomes

$$\lambda = 2\pi c\{0.872194, 0.86939, 0.5, 0.3509, 0.25\}/\omega, \quad (59)$$

c being the light velocity.

If the first or second term in Eq. (59) is in the visible spectrum, the other three terms cannot be so. For $\kappa = 0$, the principal vibration of the atom in its own plane, which has $q = \omega$, should be considered. Nicholson takes, thus, the following three wave lengths capable of appearance in the visible spectrum,

$$\lambda = 2\pi c\{1, 0.872194, 0.86939\}/\omega. \quad (60)$$

The chief nebular lines are $\lambda = 5,006.9$ Å and $\lambda = 4,363.4$ Å, the intensity of the second being about 1/10 of that of the first.\footnote{W. H. Wright, Astrophys. Journ. 16 (1902).} Assuming that the first term in Eq. (60) is the line $\lambda = 5,006.9$ Å, the second term should have a wave length $5,0069 \times 0.872194 = 4,367.0$ Å, giving the second chief line with a relative error of only 0.08%.
In the above identification of lines, one has $\omega/c = 1,255 \times 10^{-5}$, so that from Eq. (56) one gets, putting $e/mc = 1.88 \times 10^7$ and $e = 3.4 \times 10^{-10}$ (in c.g.s.),

$$a = 3.453 \times 10^{-8} \text{ cm}.$$ 

"This is the radius of the atom of nebulium, and it is in complete accord with current estimates of atomic radius, which give $10^{-8}$ as representing the order of magnitude" — so emphasizes Nicholson. He points out also that the velocity of an atomic electron relative to the light velocity $u/c = a\omega/c$ is as small as $4.33 \times 10^{-3}$, showing a sufficient margin of accuracy of neglecting magnetic effects.

About the other lines of a neutral nebulium atom (denoted by Nu), he says that though the third term in Eq. (60) should correspond to the line $\lambda = 5,006.9 \times 0.86939 = 4,352.9$ Å, this is not found in Wright's table. Nicholson notes that when a nebulium atom is ionized vibrations of higher classes become important as Schott has pointed out, and makes similar calculations including such vibrations about nebulium ions with single and double charges in the positive as well as negative senses. Investigating the correspondence of the vibrations of lower classes with the 19 lines given in Wright's table, Nicholson ascribes 3 lines to Nu, 2 lines to Nu\(^+\) (1 line possibly being due to He in part), 3 lines to Nu\(^-\) (1 line being due to H in part), 1 line to Nu\(^++\), 2 line to Nu\(^--\), and the rest lines to H or He.

In the summary of his paper, Nicholson mentions that "a connection has not been formulated between the various sets of wave length ratios". This is to say that he has not been able to derive any formula of spectral series. This is inevitable, as we have mentioned in Sec. 2 of the present chapter, for any model based on the classical mechanics without any special assumption.

In relation to this, we see his words in this paper that "some mode, by which the atom may be made of definite structure, in order to account for the fineness of spectral lines, must be invoked". Instead of doing so, Nicholson attempted at getting the radius of the nebulium atom that he took as the clue for the "elementary simple substances" he contemplated. Next to the study of the spectral lines of nebulium, he studied those of "protofluorine", which he supposed to consist of a nucleus with a positive charge of 5e and 5 electrons rotating in a ring around the nucleus.

Nicholson starts his 1911 paper "The Constitution of the Solar Corona I; Protofluorine"\(^{37}\) with the words that in his previous paper "it was shown that

there is a reason for the belief that most of the lines of the nebular spectrum
may be due to a substance which was called nebulium, a more elementary form
of matter than any yet discovered on the Earth”, and that “the theory in ques­
tion involves also the existence of other simple forms, and an adaptation of the
method of the last paper has indicated that the free vibrations of one of these
forms, which was called protofluorine when the theory was first introduced, do
not appear to be represented in the nebular spectrum”.

In these circumstances Nicholson “turns the attention, in the search for
a verification of its existence, to the coronal spectrum”. To him this seems
“natural”, because in the coronal spectrum “matter is also believed to give
evidence of its most elementary forms”. He says that “considerable evidence is
adduced to show that this substance is capable of giving rise to a large number
of the coronal lines”.

The calculations are similar to those for the case of nebulium. The neutral
atom of protofluorine is of \( \nu = 5 \) and \( n = 5 \). Corresponding to Eq. (60),
Nicholson gets

\[
\lambda = 2\pi c\{1, 0.85128, 0.86396\}/\omega. \tag{61}
\]

He identifies the first term in this equation with the fairly strong coronal line at
\( \lambda = 3,987.1 \text{ Å} \). The second and third terms should then have \( \lambda = 3,394 \text{ Å} \) and
3,445 Å respectively, in good agreement with the coronal lines at \( \lambda = 3,387.9 \text{ Å} \)
and \( \lambda = 3,454 \text{ Å} \).

This identification gives \( \omega/c = 1.576 \times 10^5 \) for the angular velocity of the
ring of the electrons. From Eq. (56) the radius of the protofluorine is therefore

\[ a = 5.802 \times 10^{-8}. \]

The protofluorine atom is a little larger than the nebulium atom. It turns out
that \( u/c = a\omega/c = 0.00917 \) brings about twice the value for nebulium atom.
The error due to neglect of magnetic effects is therefore about twice or four
times as large, according as \( u/c \) or \( (u/c) \) is the most significant order neglected.

For the protofluorine ions Nicholson makes similar analyses as for the nebu­
ilum ions. He calculates further vibrations of various classes, with methods
similar to those used in the Thomson and Nagaoka models. As a result, of the
24 known coronal lines, 14 lines are ascribed to protofluorine. That is, 3 lines
are ascribed to \( \text{Pf} \) (neutral protofluorine atom), 2 lines to each of \( \text{Pf}^\pm, \text{Pf}^{\pm\pm} \)
and \( \text{Pf}^{--} \), and 1 line to \( \text{Pf}^{++++} \). The radii of the ions are deduced to be, in
units of $10^{-8}$ cm, $a = 3.393$ for Pf$^+$, $a = 3.293$ for Pf$^-$, $a = 3.713$ for Pf$^-$ and $a = 3.499$ for Pf$^{+++}$.

§2. The quantum nature of angular momentum

In his 1911 paper "The Constitution of the Solar Corona", 38) Nicholson makes a further analysis of the spectrum of protofluorine, after saying that the nebular line at $\lambda = 4,354$ Å predicted in the analysis of the spectrum of nebulium, has been found subsequently by Wright on a plate of the Orion spectrum observed in 1908.

In the analysis of the protofluorine spectrum, Nicholson takes into consideration the quantum of action. He says "The constant of nature in terms of which these spectra can be expressed appears to be that of Planck in his recent quantum theory of energy. It is evident that the model atoms with which we deal have many of the essential characteristics of Planck’s ‘resonators’", and "If the ratio of atomic energy to principal frequency follows a regular law, the investigation will serve the double purpose of confirming the suggested origin of the spectra of astrophysics, and of giving to Planck’s theory an atomic foundation: a foundation of the kind which is now generally believed to be necessary, giving a concrete picture of the possible nature of a resonator". These words of Nicholson show clearly his substantialistic way of thinking.

Now, he takes up first the energy of the atom. If an electron rotates along a ring of radius $a$ around the nucleus of charge $\nu e$ as the center, the energy equation is

$$\frac{1}{2} m u^2 = D - \int \frac{\nu e^2}{r^2} dr = D + \frac{\nu e^2}{a},$$

where $D$ is a constant. If there are $n$ electrons in the ring, the potential energy of the system changes on account of their mutual interaction from $-n\nu e^2/a$ to

$$-\left(\frac{ne^2}{a}\right) \left(\nu - \frac{S_n}{4}\right); \quad S_n = \sum_{s=1}^{n-1} \csc\left(\frac{s\pi}{n}\right).$$

The energy equation is then given by

$$\left(\frac{1}{2}\right) m u^2 = \left(\frac{ne^2}{a}\right) \left(\nu - \frac{S_n}{4}\right) + D,$$

where $D$ means the energy in the standard configuration.

From the condition for steady motion, Eq. (56), one has

\[ ma^3 \omega^3 = e^2 \left( \nu - \frac{S_n}{4} \right) = m a u^2. \] (63)

From Eqs. (62) and (63) "the potential energy of aethereal strain" becomes

\[ D - m n a^2 \omega^2. \]

Nicholson calls \( m n a^2 \omega^2 \) simply the "energy" and investigates its value.

In the foregoing paper, the principal frequency \( q = \omega \) of the protofluorine atom is ascribed to the line at \( \lambda = 3,987.1 \) Å, so that one has

\[ \omega = \frac{2\pi c}{(3.987 \times 10^{-5})}. \]

Equation (63) then gives

\[ a^3 \omega^2 = \frac{3.6235 e^2}{m}. \]

From these two equations, one gets, using \( e = 4.7 \times 10^{-27} \) esu and \( e/mc = 1.77 \times 10^{-7} \) emu,

\[ a \omega = 3.6235 \times 10^8. \]

The ratio "energy"/frequency is then

\[ m n a^2 \omega^2 \left(\frac{2\pi}{\omega}\right) = 3.6235 \times 10^{-27}. \]

This value is about 25 times the Planck constant \( h \). Nicholson supposes that this value would be exactly 25 \( h \), if the theory becomes complete.

He interprets this result in terms of angular momentum. "Energy"/frequency is proportional to \( m n a^2 \omega \), but \( m n a^2 \omega = m n a u \) is the total angular momentum of the electrons rotating around the nucleus. "If, therefore, the constant \( h \) of Planck has, as Sommerfeld has suggested, an atomic significance, it may mean that the angular momentum of an atom can only rise or fall by discrete amounts when electrons leave or return. It is readily seen that this view presents less difficulty to the mind than the more usual interpretation, which is believed to involve an atomic constitution of energy itself". Saying so, Nicholson proposes here the quantum nature of angular momentum. It may be interesting to see that these Nicholson’s words represent a dialectic process,
in which one arrives at a new cognition in search of a thing natural in some sense to the old cognition.

Nicholson does not make any reference to Sommerfeld. But because Nicholson’s second paper on the coronal spectrum under consideration is dated April 28, 1912, he might have had some knowledge of Sommerfeld’s lecture at the 1st Solvay Conference held in autumn 1911, which we have discussed in Sec. 1, §3 of the present chapter. As we have mentioned there, Sommerfeld stressed at the 1st Solvay Conference the standpoint “to look at the existence of the molecule as a function and result of the existence of an elementary quantum of action”. In consideration of this, it is understood that Nicholson’s theory is based on the Nagaoka model but not on the Thomson model, because the size of the atom is presupposed in the latter model.

Nicholson considers now another way of expressing “energy”/frequency as the errors of the values of $e$ and $e/m$ are large. Let $\lambda$ be the wavelength of the principal line. From the relation $\omega = 2\pi c/\lambda$ and Eq. (63) he derives

$$mna^2\omega^2 \left(\frac{2\pi}{\omega}\right) = mn \left\{ \left(\frac{2\pi ce^2}{m}\right) \left(\nu - \frac{S_n}{4}\right) \right\}^{2/3} \lambda^{1/3}/c.$$ 

He thus uses

$$E = n \left(\nu - \frac{S_n}{4}\right)^{2/3} \lambda^{1/3}$$

(64)

to represent “energy”/frequency. From the result of the foregoing paper, he gets

for Pf $E = 187.04 \ (\lambda = 3918 \ \text{Å})$, 
for Pf$^+$ $E = 164.19 \ (\lambda = 4231 \ \text{Å})$, 
for Pf$^{++}$ $E = 134.29 \ (\lambda = 4586 \ \text{Å})$.

Because $E$ for Pf should corresponds to 25 in the unit of the Planck constant $h$, the equivalent Planck constant for $E$ becomes 7.482. In terms of this unit for $E$, the three values of $E$ given above are respectively, within an error smaller than 0.1%,

25, 22, 18.

Nicholson observes that “these are the first three terms of the harmonic sequence

25, 22, 18, 13, 7, 0,
which would, if it continued, give no units to the positive nucleus alone, as 
would be expected if intrinsic energy were neglected”.

“Energy”/frequency per electron, namely angular momentum per electron, 
becomes therefore, in the unit of $h$,

$$5, \frac{5}{2}, 6, \frac{6}{2}, 7$$

for the five systems ($\text{Pf}$, $\text{Pf}^+$, $\text{Pf}^{++}$, $\text{Pf}^{+++}$ and $\text{Pf}^{++++}$) under consideration 
respectively, exhibiting a regular graduation.

If this law is valid, the value of $2\pi n a^2 \omega$ for the system having $n$ electrons 
and the principal frequency $f$ would be

$$\frac{1}{2} (15 - n) n f h.$$ 

(65)

Thus, Nicholson concludes that it becomes possible to express the whole 
spectrum of protofluorine in the form of a series, or a set of series, depending on 
the natural numbers, though the series is not of the types of Balmer or Rydberg. 
He adds, however, that these series relate to elements whose structure is more 
complicated than a single ring.

In his next paper “The Constitution of the Solar Corona III”, Nicholson 
investigates this problem and states the following: “The quantum theory has 
apparently not been put forward as an explanation of ‘series’ spectra, consisting 
of a large number of related lines given by a comparatively simple atom. Yet, 
in the belief of the writer, it furnishes the true explanation in certain cases, 
and we are led to suppose that lines of a series may not emanate from the 
same atom, but atoms whose internal angular momenta have, by radiation or 
otherwise, run down by various discrete amounts from a standard value.”

His subsequent words are very noteworthy: “For example, in this view 
there are various kinds of hydrogen atom, identical in chemical properties and 
even in weight, but different in their internal motions. ⋯ If an atom loses its 
energy by definite amounts, instead of in a continuous manner, it should show 
a series spectrum with lines corresponding to each of the stages. Moreover, its 
incapacity for radiating in a continuous way secures sharpness of the lines.”

This Nicholson’s thinking is, as we shall see in Sec. 4, § 2 below, fairly close 
to the fundamental assumptions in the Bohr model of the atom. Nicholson’s 
thinking includes, however, a difficulty to be overcome by Bohr. That is, 

Nicholson's thinking can be said to make an important step toward Bohr's theory.

Now, about the protofluorine atom with three electrons missing (Pf+++) for which \( n = 2 \) and \( \nu = 5 \), Nicholson says that a spectral series can be found in the following way. From Eq. (64) one has

\[
\lambda = \frac{(\frac{1}{2}E)^3}{(\nu - \frac{1}{4})^2}.
\]  

Thus, at any stage in the radiation of the charged atom, the number of quanta present is proportional to the cubic root of the emitted wave number. Accordingly, the series is such that \( \lambda^{1/3} \) differs by a constant quantity. The principal line of this system has been found to be \( \lambda = 5,072 \) Å, and should be the "head" line of the series. This line corresponds to \( E = 97.107 \). Decreasing \( E \) by successive amounts of 2.2446, one gets a series of lines

\[
\lambda = (48.553 - 1.1223r)^{3/4}/(4.75)^2; \quad r = 0, 1, 2, \ldots,
\]  

that is, they are numerically given by

<table>
<thead>
<tr>
<th>( r )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda )</td>
<td>5073</td>
<td>4725</td>
<td>4400</td>
<td>4086.5</td>
<td>3788</td>
<td>3506</td>
<td>3238</td>
</tr>
<tr>
<td>( \lambda' )</td>
<td>5073</td>
<td>4275</td>
<td>4400</td>
<td>4087.5</td>
<td>⋯</td>
<td>3606</td>
<td>⋯</td>
</tr>
</tbody>
</table>

In this table \( \lambda' \) denotes the values observed by Dyson.\(^{40}\)

The wavelengths of the principal lines for the other values of \( n \) are shown in the following table, together with that for \( n = 2 \) given above,

\[
\begin{align*}
 n = 5, & \quad \lambda = 3987; \quad n = 4, & \quad \lambda = 4231 \\
 n = 3, & \quad \lambda = 4586; \quad n = 2, & \quad \lambda = 5073
\end{align*}
\]

Each of these wavelengths represents the "head" of the respective series of lines. Let \( Q_n \) be defined as

\[
Q_n = E/n = \lambda^{1/3}(\nu - S_n/4)^{2/3}/0.06235.
\]  

\(^{40}\)Dyson, Phil. Trans. A26 (1906), 451.
$Q_n$ is the number of quanta per electron evaluated in the unit of $\hbar/5!$. For the "heads" the values of $Q_n$ are

$$Q_5 = 600, \quad Q_4 = 658\frac{1}{2}, \quad Q_3 = 718, \quad Q_2 = 778\frac{1}{2}.$$  

Their differences $58\frac{1}{2}, 59\frac{1}{2}, 60\frac{1}{2}$ are accurately harmonic, so that the law $E/n = A + Bn$ is satisfied.

It follows that for the atom with one electron, the principal line should have $Q_1 = 840$, corresponding to the wavelength

$$\lambda = (52.374)^2/25 = 5746.5 \text{ Å}.$$  

This is outside the observed range. For the atom with two electrons (Pf+++), because the numerical coefficient in Eq. (67), 1.1223, corresponds exactly to 18 quanta, the values of $Q_2$ for the series of lines given in Eq. (67) can be written as

$$Q_2 = 778\frac{1}{2} - 18r; \quad r = 0, 1, 2, \ldots.$$  

(69)

In this way, paying attention to Sommerfeld's suggestion of the importance of the quantum of action in molecular processes, Nicholson arrived at the concept of the quantum nature of angular momentum, and attempted on this basis to understand spectral series. Even if the atoms he treated were those of nebulium or protofluorine, which should be the "elementary simple substances" in existence in nebula or solar corona,41) the spectral series he obtained were not in accordance with Ritz's combination rule.

In August 1912, four months after Nicholson wrote the paper "The Constitution of the Solar Corona II", in which Nicholson proposed the quantum nature of angular momentum, Bohr wrote the paper "On the Theory of the Decrease of Velocity of Moving Electrified Particles on Passing Through Matter".42) As we have mentioned in Vol. I, in this paper Bohr confirmed that the number of electrons in the hydrogen atom is just one. By reason of this Nicholson's theory of spectral series, in which an equilibrium among neutral atoms and positively as well as negatively charged atoms was presumed, turned out to

41) I. S. Bowen showed in 1927 [Nature 120, 473] that the lines of nebulium discussed by Nicholson are due to nitrogen ion N+++ and oxygen ions O++ and O+++.

42) N. Bohr, Phil. Mag. 25 (1913), 10.
be ineffective. However, Nicholson's attempt to introduce the quantum nature of angular momentum remained effective in the Bohr model of the atom.\footnote{About inadequate valuations of Nicholson's work by some historians of science, confer Footnote 71} in Sec. 4, \S 6 of the present chapter.

4. The Bohr Model of the Atom

\S 1. Characteristic X-rays and frequencies of atomic electrons

In the paper just cited above, Bohr, besides confirming the number of electrons in the hydrogen atom, made an estimate of the frequencies of the characteristic X-rays of oxygen, and compared them with those of the atomic electrons of oxygen with good result.

As we have mentioned in Sec. 6, \S 5, Chap. 2, Vol. I, Bohr puts forth the fundamental basis of his theory of velocity loss upon the vibration of atomic electrons. A charged particle passing through matter gives impacts to atomic electrons. If the vibrational periods of atomic electrons are of the same order of magnitude as the collision time, the motions of atomic electrons are changed during the collision, and the charged particle losses its energy so that its velocity diminishes. The velocity loss of a charged particle is defined as the rate of the decreases of its velocity per unit distance that it passes through matter.

Let \( M, E, \) and \( V \) be the mass, energy and velocity of the charged particle, and \( m \) and \( e \) be the mass of an electron and the elementary electric charge, respectively. If the charged particle changes its velocity by \( dV \) in its passage through \( dx \), the velocity loss is given by\footnote{Cf. p. 227, Sec. 6, \S 5, Chap. 2, Vol. I.}

\[
\frac{dV}{dx} = -\frac{4\pi e^2 E^2 N}{mMV^2} \sum_{s=1}^{r} \log \left( \frac{V^3 k m M}{n_s e E (m + M)} \right).
\]

In this equation, \( k = 1.123 \) is a constant arising from integration over the impact parameter of the charged particle, \( n_s \) is the frequency of the \( s \)th atomic electron, and \( N \) is the number of atoms in the unit volume, each atom being supposed to have \( r \) electrons.

The velocity loss of \( \alpha \)-rays in oxygen is, from the observed values by Adams,\footnote{E. P. Adams, Phys. Rev. 24 (1907), 113.} \( dV/dx = -2.07 \times 10^8 \) at \( V = 1.35 \times 10^9 \) and \( dV/dx = -1.24 \times 10^8 \)
at $V = 1.75 \times 10^9$ in c.g.s., so that one has from the above equation

$$\sum_{s=1}^{r} \{\log(n_s \times 10^{-19}) + 0.59\} = -47,$$

$$\sum_{s=1}^{r} \{\log(n_s \times 10^{-19}) - 0.18\} = -61.$$ 

Taking the difference between them, one gets $0.77r = 14$, or $r = 18$. Bohr considers this result very satisfactory, because according to Rutherford's theory an oxygen atom should have sixteen electrons.

It is known from the dispersion of light that an oxygen molecule possesses four electrons of frequency $n_s = 2.25 \times 10^{16}$, so that

$$\sum_{s=1}^{r} \log(n_s \times 10^{-19}) = -34.$$ 

If the other twelve electrons are assumed to have an identical frequency $n'$, one gets $n' = 0.5 \times 10^{18}$ from this equation.

Although about the higher frequencies of oxygen very little is known, Bohr thinks that they can be inferred from experiments on characteristic X-rays. Whiddington\(^{46}\) has found that the velocity of an electron just sufficient to excite the characteristic X-rays in an element of mass number $A$ is equal to $A \times 10^8$ cm/s. Bohr considers then that "according to Planck’s theory of radiation we further have that the smallest quantity of energy which can be radiated out from an atomic vibrator is equal to $\nu k$, where $\nu$ is the number of vibrations per second and $k = 6.55 \times 10^{-27}$,\(^{47}\) and "this quantity must be expected to be equal to, or at least of the same order of magnitude as, the kinetic energy of an electron of velocity just sufficient to excite the radiation".

Thus, he gets $\nu = 6.7A^2 \times 10^{14}$ by putting $\nu k = (m/2)A^2 \times 10^{16}$. Assuming that Whiddington’s rule applies well to the case of oxygen, for which there is no experiment on characteristic X-rays, Bohr puts $A = 16$ and derives $\nu = 1.7 \times 10^{17}$ and therefore $2\pi\nu = 1.1 \times 10^{18}$. He says that it is remarkable that this value agrees with the value of $n'$ obtained above in order of magnitude.

In this way, Bohr supposed an electron oscillating in an atom to be Planck’s oscillator, and related the quantum nature of its energy with the energy


\(^{47}\)The constant $k$ used here by Bohr is of course equal to the Planck constant.
required of an incident electron to excite characteristic X-rays. As we have already seen, Einstein studied the specific heat of solids on the basis of Planck's theory, by applying the quantum nature of the energy of resonator to the oscillation of atoms around their equilibrium points. Haas tried to relate the oscillation of an electron in the positively charged sphere of the Thomson model to Planck's theory, but the quantum nature of the energy of resonator was not used because of his introduction of the maximum temperature. Sommerfeld treated the photoelectric effect by supposing an atomic electron making simple harmonic motion, but the quantum nature of the energy of the resonator was not taken up directly, because his theory was based on the hypothesis of quantum of action in which importance was attached to the action integral.

In the case of Nicholson's theory discussed in the preceding section, he used the quantum nature of the energy of the resonator to compare the quantity energy/frequency with \( h \), saying that "the model of atom dealt with have many of the essential characteristics of Planck's resonators", and arrived at the concept of quantum of angular momentum. In Bohr's paper under consideration, the substantial structure of the hydrogen atom was confirmed. In the Bohr model of the atom, the quantum concept was applied to the hydrogen atom, which has a very simple structure, bringing about a great success.

§2. The quantum of action and the fundamental assumptions in the Bohr model

In a series of papers "On the Constitution of Atoms and Molecules", Part I,\(^{48}\) Part II\(^{49}\) and Part III,\(^{50}\) Bohr investigated the spectrum and structure of hydrogen atom, and furthermore those of heavier atoms and molecules. The essence of Bohr's theory is concentrated in the case of the hydrogen atom. In many text books of general physics or atomic physics, descriptions are given of the Bohr model mixed with the old quantum mechanics. In the present book analyses will be made of the Bohr model with our emphasis put on the development of logic in the model itself.

In the introduction to the series, Bohr notes that the forces acting on the electrons in the Thomson model allow of certain configurations and motions of the electrons for which the system is in a stable equilibrium, but

\(^{48}\)N. Bohr, Phil. Mag. 26 (1913), 1.
\(^{49}\)N. Bohr, Phil. Mag. 26 (1913), 476.
\(^{50}\)N. Bohr, Phil. Mag. 26 (1913), 857.
such configurations apparently do not exist in the Rutherford model. He then mentions that "among the quantities characterizing the first atom a quantity appears — radius of the positive sphere — of dimensions of a length and of the same order of magnitude as the linear extension of the atom, while such a length does not appear among the quantities characterizing the second atom, viz. the charges and masses of the electrons and the positive nucleus; nor can it be determined solely by the help of the latter quantities".

As we have seen in Sec. 1, §2 of the present chapter, Haas attempted to derive the quantum of action $h$ from the radius of the positive sphere $a$, the elementary electric charge $e$ and the mass of electron $m$. About this, Sommerfeld stressed at the 1st Solvay Conference, as we have also seen in Sec. 1, §3, the "standpoint not to explain $h$ from the molecular dimensions, but to look at the existence of the molecule as a function and result of the existence of an elementary quantum of action". Since Rutherford confirmed the Nagaoka model that was in opposition to the Thomson model, and Bohr has concluded that each hydrogen atom has only one electron, this idea of Sommerfeld has come to have a more important meaning.

Bohr himself cites the Proceedings of the 1st Solvay Conference, in a footnote in his paper under consideration, as “See f. inst., 'Theorie du ravonnement et les quanta, Rapports de la reunion à Bruxelles, Nov. 1911. Paris, 1912", to note "a general acknowledgment of the inadequacy of the classical electrodynamics in describing the behaviour of systems of atomic size". He continues to say as follows. "Whatever the alternation in the laws of motion of the electrons may be, it seems necessary to introduce in the laws in question a quantity foreign to the classical electrodynamics, i.e. Planck's constant, or, as it often is called, the elementary quantum of action. By the introduction of this quantity the question of the stable configuration of the electrons in the atoms is essentially changed, as this constant is of such dimensions and magnitude that it, together with the mass and charge of the particles, can determine a length of the order of magnitude required".$^{51}$

Now in "Part I. — Binding of Electrons by Positive Nuclei", Bohr considers a simple system consisting of a positively charged nucleus of very small size and an electron describing closed orbits around it. If it is assumed that there is no radiation of energy the electron will describe stationary elliptical orbits. Let $\omega$ be the frequency of revolution, $2a$ the major-axis of the orbit, and $W$

---

$^{51}$He returns to this point in Part I (cf. p. 85 of the present section).
the amount of energy which must be transferred to the system in order to remove the electron to an finitely great distance from the nucleus. From the Newtonian mechanics for the Kepler motion, we get

\[ \omega = \frac{\sqrt{2}}{\pi} \frac{W^{3/2}}{eE\sqrt{m}}, \quad 2a = \frac{eE}{W}, \]  

(70)

where \(-e\) and \(E\) are the charges of the electron and nucleus, respectively, and \(m\) is the mass of the electron. If the value of \(W\) is not given, there will be no values of \(w\) and \(a\) characteristic of the system in question.

If we take into account effect of the energy radiation due to the acceleration of the electron, it will approach the nucleus, describing orbits of smaller and smaller dimensions, and with greater and greater frequency. The process will go on until the dimensions of the orbit are of the same order of magnitude as that of the nucleus. Bohr thinks, however, that in the first place the actual atoms in their "permanent state" seem to have absolutely fixed dimensions and frequencies, and in the second place in any molecular process after a certain amount of energy characteristic of the systems is radiated out, the systems will again settle down in a state of equilibrium.

So, taking the case in which the electron at the beginning was at a great distance from the nucleus, Bohr assumes that the electron after the interaction has taken place settles down in a stationary orbit around the nucleus. The orbit is assumed to be circular. This assumption will make no alternation in the calculations for systems containing only a single electron.

The frequency of the electron at the beginning is zero, and in the final orbit is \(\omega\). Bohr assumes that during the binding of the electron, a radiation of a frequency equal to their average \(\omega/2\) is emitted. According to Plank's theory, he puts

\[ W = \tau h\omega/2, \]  

(71)

where \(\tau\) is an integer. With the help of Eq. (70), he gets

\[ W = \frac{2\pi^2 me^2 E^2}{\tau^2 h^2}, \quad \omega = \frac{4\pi^2 me^2 E^2}{\tau^3 h^3}, \quad 2a = \frac{\pi^3 h^2}{2\pi^3 meE}. \]  

(72)

If in these expressions different values are given to \(\tau\) a series of values of \(W, \omega\) and \(a\) is obtained, "corresponding to a series of configurations of the system". Bohr assumed that "these configurations will correspond to states in which there is no radiation of energy". Supposing that these states "consequently
The Bohr Model of the Atom

will be stationary as long as the system is not disturbed from the outside”, Bohr assumes here the existence of a series of stationary states in the atom.

Because $W$ is greatest for $\tau = 1$, $\tau = 1$ corresponds to the most stable state of the system. For $\tau = 1$ and $E = e$, with the use of the values in c.g.s.,

$$e = 4.7 \times 10^{-10}, \quad e/m = 5.3 \times 10^{17} \quad h = 6.5 \times 10^{-27},$$

Bohr gets

$$2a = 1.1 \times 10^{-8} \text{ cm}, \quad \omega = 6.2 \times 10^{15} \text{ s}^{-1}, \quad W/e = 13 \text{ V}.$$

These values are of the same order of magnitude, respectively, as the linear dimensions, optical frequencies and ionization potentials of atoms.

Bohr replies in this way to Sommerfeld’s suggestion mentioned above that “the existence of molecules is the result and function of the existence of the quantum of action”. Bohr writes the following. “The general importance of Planck’s theory for the discussion of the behaviour of atomic systems was originally pointed out by Einstein. The considerations of Einstein have been developed and applied on a number of different phenomena, especially by Stark, Nernst, and Sommerfeld. The agreement as to the order of magnitude between values observed for the frequencies and dimensions of the atoms, and values for these quantities calculated by considerations similar to those given above, has been the subject of much discussion. It was first pointed out by Haas, in an attempt to explain the meaning and value of Planck’s constant on the basis of J. J Thomson’s atom-model, by the help of the linear dimensions and frequency of a hydrogen atom”. One will clearly see the meaning of these Bohr’s words from what we have discussed in Vol. I as well as in the foregoing and present chapters in the present volume.

Bohr takes then Nicholson’s works, which we have discussed in the last section, and says that “Nicholson has obtained a relation to Planck’s theory showing that the ratios between the wave-length of different sets of lines of the coronal spectrum can be accounted for with great accuracy by assuming that the ratio between the energy of the system and the frequency of rotation of the ring is equal to an entire multiple of Planck’s constant”.

In Nicholson’s theory there have been, however, certain difficulties to be overcome. Bohr analyzes them in the following way. “In Nicholson’s calculations the frequency of lines in a line-spectrum is identified with the frequency of vibration of a mechanical system in a distinctly indicated state of equilibrium.
As a relation from Planck’s theory is used, we might expect that the radiation is sent out in quanta; but systems like those considered, in which the frequency is a function of the energy, cannot emit a finite amount of a homogeneous radiation; for, as soon as the emission of radiation is started, the energy and also the frequency of the system are altered.”\(^{52}\) Besides this essential difficulty, Bohr points out that “according to the calculation of Nicholson, the systems are unstable for some mode of vibration”, and “the theory in the form given does not seem to be able to account for the well-known law’s of Balmer and Rydberg”.

Nicholson attempted to avoid these difficulties, as we have described in the preceding section, by supposing that in nebulae or coronae, under the condition of small radiative rates, neutral atoms of an element with four or five electrons and positive and negative ions of the respective element would transform into one another, building an equilibrium in which they could only be in states determined by the quantum nature of angular momentum. For the hydrogen element whose atom has only one electron, however, Nicholson’s supposition could not be applied. The cognition of the new substance, that is, the confirmation of the hydrogen atom having only one electron, has forced a theory definitely different from the previous theories to start.

Thus, Bohr restates briefly the idea underlying Eq. (72) and the assumptions to deduce his results, as follows — “The principal assumptions used are:

1. That the dynamical equilibrium of the systems in the stationary states can be treated by the help of the ordinary mechanics, while the passing of the systems between different stationary states cannot be treated on that basis.\(^{53}\)

2. That the latter process is followed by the emission of a *homogeneous* radiation, for which the relation between the frequency and the amount of energy emitted is that given by Planck’s theory.”

Instead of Nicholson’s supposition, Bohr makes the assumption that in a single atom there exist stationary states which can be treated by the ordinary mechanics. Also Bohr attributes the source of the energy of light to the energy

\(^{52}\)Bohr calls a radiation of a definite frequency (a monochromatic light) “a homogeneous radiation”.

\(^{53}\)Bohr uses here the word “passing”, but not yet the word “transition” (cf. p. 143, Sec. 2, §7, Chap. 3 of the present volume.)
difference between stationery states of an atom, while Nicholson has attributed it to a group of atoms and ions.

Bohr writes that "the first assumption seems to present itself, for it is known that the ordinary mechanics cannot have an absolute validity, but will only hold in calculations of certain mean values of the motion of the electrons", and "the second assumption is in obvious contrast to the ordinary ideas of electrodynamics, but appears to be necessary in order to account for the experimental facts".

The treatment of the dynamical equilibrium in the stationary states by the ordinary mechanics can be said to be a substantialistic step. On the other hand, the mechanism of the passing of the system between stationary states should be treated phenomenologically. That is, it can be said that Bohr made a clear-cut separation between substantialistic and phenomenological steps.

§3. The spectral formula and the angular momentum

Bohr now intends to relate Eq. (72) to Balmer's spectral formula. Citing his paper which we have quoted in §1 of the present section, he notes that "general evidence indicates that an atom of hydrogen consists simply of a single electron rotating round a positive nucleus of charge e". Putting therefore \( E = e \) in Eq. (72), he gets for the total amount of energy radiated out by the formation of one of the stationary states,

\[
W_{\tau} = \frac{2\pi^2me^4}{h^2\tau^2}.
\]  

(73)

The energy emitted by the passing of the system from a state corresponding to \( \tau = \tau_1 \) that corresponding to \( \tau = \tau_2 \) is consequently

\[
W_{\tau_2} - W_{\tau_1} = \frac{2\pi^2me^4}{h^2} \left( \frac{1}{\tau_2^2} - \frac{1}{\tau_1^2} \right).
\]  

(74)

Bohr introduces here a new assumption. That is, he says that "if we now suppose that the radiation in question is homogeneous, and that the amount of energy emitted is equal to \( h\nu \), where \( \nu \) is the frequency of the radiation, we get

\[
W_{\tau_2} - W_{\tau_1} = h\nu.
\]  

(75)

As we have mentioned above, in the derivation of Eq. (72) Bohr assumes that when the electron passes from a state in which it is very distant from the
nucleus and has almost no velocity to a stationary state in which it rotates with frequency $\omega$ around the nucleus, a radiation with the frequency $\omega / 2$ equal to the average of the rotational frequencies in these states is emitted, and the energy of the radiation is given by $\tau h \omega / 2$. If one follows this Bohr’s assumption as it is, one would have $(\omega_1 + \omega_2) / 2$ for the frequency of the radiation in question, where $\omega_1$ and $\omega_2$ denote the rotational frequencies in the states corresponding to $\tau = \tau_1$ and $\tau = \tau_2$, respectively, so that one should have, instead of Eq. (75),

$$W_{\tau_2} - W_{\tau_1} = \tau h \frac{(\omega_2 + \omega_1)}{2}.$$ 

If the state corresponding to $\tau = \tau_1$ is taken to be the state before the binding for which $W_{\tau_1} = 0$ and $\omega_1 = 0$, this equation reproduces Eq. (71), but Eq. (75) does not. Bohr’s new assumption is inconsistent with the assumption made for Eq. (71) which gives $W$.

The purpose for making this new assumption is to get a spectral formula of Balmer–Rydberg’s type. That is, Bohr derives from Eqs. (74) and (75)

$$\nu = \frac{2\pi^2 me^4}{\hbar^3} \left( \frac{1}{\tau_2^2} - \frac{1}{\tau_1^2} \right).$$

This expression agrees with the observed spectral series in “quantitative as well as qualitative” senses. The values of $e$, $e/m$ and $h$ used above in §2 give

$$\frac{2\pi^2 me^4}{\hbar^3} = 3.1 \times 10^{15}. $$

Bohr says that this theoretical value agrees with the observed value $3.290 \times 10^{15}$ “inside the uncertainty due to experimental errors in the constants entering in the expression”. As we shall see shortly about Eq. (76), Bohr reinterprets Eq. (71) on account of the inconsistency mentioned just above.\(^{54}\)

Now, Eqs. (74) and (75) with $\tau_2 = 2$ give the ordinary Balmer series, and that with $\tau_2 = 3$ the series in the infrared observed by Paschen\(^{55}\) and previously predicted by Ritz. Bohr remarks that the fact that it has not been possible to observe more than 12 lines of the Balmer series in experiments


\(^{55}\)F. Paschen, Ann. d. Phys. 27 (1908), 565.
with vacuum tubes, while 33 lines are observed in the spectra of some celestial bodies is just what we should expect from his theory. According to Eq. (72), for \( \tau = 12 \) the diameter of the orbit of the electron is \( 1.6 \times 10^{-6} \) cm, equal to the mean distance between the molecules in a gas at a pressure of 7 mmHg. For \( \tau = 33 \) it is \( 1.2 \times 10^{-5} \) cm, corresponding to the mean distance at a pressure of 0.02 mmHg. Bohr remarks that it might be possible to observe lines corresponding to higher \( \tau \)'s, by investigation of the absorption spectrum of the gas.

For the helium atom, considering the binding of a single electron, Bohr replaces \( e^2 \) with \( Ee \) and puts \( E = 2e \) in the equation before the last to get

\[
\nu = \frac{2\pi^2 me^4}{h^3} \left\{ \frac{1}{(\tau_2/2)^2} - \frac{1}{(\tau_1/2)^2} \right\}.
\]

This formula with \( \tau_2 = 3 \) gives two of the series observed by Fowler.\(^{56}\) The formula with \( \tau_2 = 4 \) gives the series observed by Pickering\(^{57}\) in the spectrum of \( \zeta \) Puppis. Every second of the lines in this series is identical with a line in the Balmer series of the hydrogen atom. From the fact that these lines are of a greater intensity than the rest of the lines in the series, Bohr infers the existence of the hydrogen in this star.

According to Rydberg's theory with the generalization due to Ritz,\(^{58}\) the frequency of a spectral line of an element are given by

\[
\nu = F_r(\tau_1) - F_s(\tau_2),
\]

where \( \tau_1 \) and \( \tau_2 \) are integers, and \( F_1, F_2, F_3, \ldots \) are functions of \( \tau \) which are approximately equal to \( K/(\tau + a_1)^2 \), \( K/(\tau + a_2)^2 \), \( K/(\tau + a_3)^2 \), \ldots \( K \) being a universal constant.

Bohr argues that the universality of the constant \( K \) "can be simply explained". That is, according to him, the circumstance that the frequency can be written as a difference between two functions of an integer suggests an origin of the lines similar to the one we have assumed for hydrogen. On the assumption that the spectrum in question corresponds to radiation emitted during the binding of an electron and the whole system is neutral, the force acting on the electron at a great distance from the nucleus and the other electrons previously bound will be very nearly the same as in the case of the binding of


\(^{58}\) W. Ritz, Phys. Zeit, 9 (1908), 521 [the same as Ref. 31] on p. 63 in the present chapter.
an electron to a hydrogen nucleus. The energy corresponding to a stationary state will therefore, for large \( \tau \), be very nearly equal to that given by Eq. (72) with \( E = e \). For large \( \tau \) great it consequently results that

\[
\lim(\tau^2 F_1(\tau)) = \lim(\tau^2 F_2(\tau)) = \cdots = \frac{2\pi^2 me^4}{h^3},
\]

in conformity with Rydberg's theory.

Then, Bohr returns to the problem of the inconsistency between Eq. (71) and Eq. (75) which we have remarked above. About Eq. (71), he has "assumed that the different stationary states correspond to an emission of a different number of energy-quanta". He says, however, that "this assumption may be regarded as improbable, for as soon as one quantum is sent out the frequency is altered". From the very reason that as soon as one quantum is sent out the frequency is altered, he has used the average frequency \( \omega/2 \), but with this he could not derive, as we have mentioned above, the spectral formula. Thus, Bohr says that "it has not been necessary, in order to account for the law of the spectra by the help of the expressions (72) for the stationary states, to assume that in any case a radiation is sent out corresponding to more than a single energy quantum \( h\nu \)," \(^59\) and makes the following formulation.

In other words, he assumes that "the ratio between the total amount of energy emitted and frequency of revolution of the electron for the different stationary states is given by

\[
W = f(\tau)h\nu,
\]

instead of Eq. (71)." \(^60\) \( W \) is originally "the energy which must be transferred to the system in order to remove the electron to an infinitely great distance apart from the nucleus", and corresponds to one-half of Nicholson's 'energy' as Bohr himself indicates. Equation (76) assumed by Bohr derives, therefore, from the same idea as Nicholson's in which the 'energy'/frequency of the protofluorine atom has been seen to be some multiple of \( h \). In Eq. (71) \( \tau \) means the number of the energy quanta of the emitted radiation, while in Eq. (76) \( f(\tau) \) is some number to determine the stationary states of the system, similar to Nicholson's \((15 - n)/2 \) in Eq. (65).

---

\(^{59}\) Equation (72) corresponds to Eq. (3) in Bohr's paper.

\(^{60}\) Equation (71) corresponds to Eq. (2) in Bohr's paper.
Differently from Nicholson, however, Bohr seeks such \( f(\tau) \) that gives the Balmer formula. That is, from Eqs. (70) and (76) we have

\[
W = \frac{\pi^2 me^2 E^2}{2^3 f^2(\tau)} , \quad \omega = \frac{\pi^2 me^2 E^2}{2^3 f^2(\tau)} ,
\]

and with the use of Eq. (75)

\[
\nu = \frac{\pi^2 me^2 E^2}{2h^3} \left\{ \frac{1}{f^2(\tau_2)} - \frac{1}{f^2(\tau_1)} \right\} . \tag{77}
\]

"We see that in order to get an expression of the same form as the Balmer series we must put \( f(\tau) = c \tau \)." Bohr shows the proportional constant \( c \) is equal to 1/2, by an interesting way of thinking that should be said the germ of the correspondence principle, which will be taken up in the next chapter.

For the passing of the system between the stationary states corresponding to \( \tau = N \) and \( \tau = N - 1 \), we have, with \( f(\tau) = c \tau \),

\[
\nu = \frac{\pi^2 me^2 E^2}{2c^3 h^3} \left[ \frac{2N - 1}{N^2(N - 1)^2} \right] .
\]

On the other hand, the angular frequency in these states are given by

\[
\omega_N = \frac{\pi^2 me^2 E^2}{2c^3 h^3 N^3} , \quad \omega_{N-1} = \frac{\pi^2 me^2 E^2}{2c^3 h^3 (N - 1)^3} .
\]

If \( N \gg 1 \), we have \( \omega_N / \omega_{N-1} \approx 1 \). Bohr considers that "according to the ordinary electrodynamics we should therefore expect that the frequency of revolution also is very nearly equal to 1". From the relation \( \nu \approx 2c\omega_N \) for \( N \gg 1 \), this condition will be satisfied only if \( c = 1/2 \). Equation (76) will then be identical with Eq. (73).

For the passing of the system between the states corresponding to \( \tau = N \) and \( \tau = N - n(N \gg n) \), we have in the same approximation

\[
\nu = n\omega . \tag{78}
\]

From this relation, Bohr argues that "the possibility of emission of a radiation of such a frequency may also be interpreted by analogy with the ordinary electrodynamics, as an electron rotating round the nucleus in an elliptical orbit will emit radiation which according to Fourier's theorem can be resolved into homogeneous components, the frequencies of which are \( n\omega \), if \( \omega \) is the frequency of resolution of the electron". In these Bohr's words we see the germ
of the correspondence principle. To arrive at the correspondence principle, some theory is needed concerning the intensity of radiation, besides the theory of the frequency of radiation. It is furnished by the concept of the transition probability due to Einstein, as we shall see in the next chapter.\(^{61}\)

Now, Bohr reinterprets Eq. (71). He writes that “we are thus led to assume that the interpretation of the Eq. (71) is not that the different numbers of energy-quanta, but that the frequency of the energy emitted during the passing of the system from a state in which no energy is yet radiated out to one of the different stationary states, is equal to different multiples of \(\omega/2\), where \(\omega\) is the frequency of revolution of the electron in the state considered”.\(^{62}\) In short, he sees Eq. (71) in the form \(W = h \times (\pi \omega/2)\), and supposes that the number of the energy quantum emitted is 1 as in Eq. (75).

To Eq. (71), however, “it is possible to give a very simple interpretation by help of symbols taken from the ordinary mechanics”. Denoting the angular momentum and kinetic energy of the electron round the nucleus by \(M\) and \(T\) respectively, we have \(\pi M = T/\omega\). For a circular orbit we further have \(T = W\). From Eq. (71), we therefore get

\[
M = \pi M_0, \tag{79}
\]

where

\[
M_0 = \frac{\hbar}{2\pi} = 1.04 \times 10^{-27}.
\]

Thus, together with the principal assumptions, Eqs. (75) and (79) form the core of the Bohr model. Eq. (79) is concerned with the atom, and does not include any assumption about radiation as Eq. (71) does. The connection with radiation is now given by Eq. (75).

Bohr states that “the possible importance of the angular momentum in the discussion of atomic systems in relation to Planck’s theory is emphasized by Nicholson”. In fact, as we have seen in Sec. 3, §2 of the present chapter, Nicholson arrived at the concept of the quantum nature of angular momentum by taking notice of the quantity energy/frequency that is equivalent to the

\(^{61}\)Cf. Sec. 4, Chap. 3. Bohr considered in 1918 that for \(N \gg n\) not only the frequency of radiation is given by the Fourier frequency \(n\omega\), but also the transition probability to determine the intensity of radiation is given approximately by the Fourier amplitude of the electric dipole moment in the classical electromagnetics, and named this idea the correspondence principle.

\(^{62}\)Equation (71) corresponds to Eq. (2) in Bohr’s paper.
angular momentum. It is needless to say anew that, in the development of science and technology, the grasp of the existence of a problem, the importance of things, the direction of solution, and so on, plays important roles.

§4. The permanent state and the angular momentum

Bohr calls the state during whose formation the greatest amount of energy is emitted the "permanent" state among the stationary states. According to Eq. (72), this state corresponds to $\tau = 1$. Bohr considers that atoms on ordinary conditions are in their respective permanent states. Making use of the concept of stationary and permanent states, he studies the absorption of radiation by atom, the photoelectric effect and the emission of X-rays. He further attempts to determine the permanent states of atoms having many electrons, in terms of "the constancy of the angular momentum".

From the circumstance that certain kinds of substance in non-luminous state, such as sodium vapour for example, absorbs radiation corresponding to lines in the spectrum of the substance, Bohr concludes that the lines in question are emitted during the passing of the system between two states, one of which is the permanent state.

We may expect also the possibility of absorption of radiation corresponding to the passing between a stationary state and a state in which the electron is free. Therefore, "we get the same expression for the kinetic energy of an electron ejected from an atom by photoelectric effect as that deduced by Einstein,\(^{63}\) i.e. $T = hv - W$". Further, Bohr argues that "we may assume that the homogeneous Röntgen radiation is emitted during the settling down of the system after one of the firmly bound electrons escapes, e.g. by impact of cathode particles".\(^{64}\) About this problem, he makes a detailed calculation in Part II.

According to Bohr, the permanent state of an atomic system is determined, if the system contains one electron, by "the condition that the angular momentum of the electron round the nucleus is equal to $\hbar/2\pi$". For an atomic system in which more electrons are bound by a nucleus, "a configuration of the electrons which presents itself as a permanent state is one in which the electrons are arranged in a ring around the nucleus".

If $n$ rotating electrons are arranged at equal angular intervals around the circumference of a circle of radius $a$, from the balance of the centrifugal and


\(^{64}\)J. J. Thomson, Phil. Mag. 23 (1912), 456.
The Quantum of Action and Atomic Models

centripetal forces we have
\[
\frac{2T}{a} = \left( \frac{e}{a^2} \right) (E - es_n). \tag{80}
\]
Bohr's \( s_n \) in this equation is equal to \( 1/4 \) of Nicholson's \( S_n \), so that this equation corresponds to Eq. (56). From Eq. (80) we get for the angular frequency
\[
\omega = \frac{1}{2\pi} \sqrt{\frac{e(E - es_n)}{ma^3}}. \tag{81}
\]

The total energy \( W \) to be transferred to the system, in order to remove all the electrons to infinite distance from the nucleus and from each other, is
\[
W = -P - nT = \left( \frac{ne}{2a} \right) (E - es_n) = nT, \tag{82}
\]
being equal to the total kinetic energy of the electrons. In this equation, \( P \) denotes the potential energy of the system and is given by
\[
P = - \left( \frac{ne}{2a} \right) (E - es_n). \tag{83}
\]
The balance equation for the case of \( n \) electrons rotating along a single ring is therefore obtained from that for the case of an rotating electron, by replacing \( E \) and \( W \) respectively with \( E - es_n \) and \( W/n \). On account of this, Bohr assumes that in the permanent state of the system of \( n \) electrons, the angular momentum of each electron is equal to \( h/2\pi \).

"As Nicholson's calculations show, the answer to the question of stability differs very much in the two cases." For displacements of the electrons perpendicular to the plane of the ring, the ring is in general stable, if the number of electrons is not great. On the other, for displacements of the electrons in the plane of the ring, the ring is unstable in the cases considered by Nicholson. Bohr considers, however, that "according to the point of view taken in this paper, the question of stability for displacements of the electrons in the plane of the ring is most intimately connected with the question of the binding of the electrons, and like the latter cannot be treated on the basis of the ordinary mechanics". Bohr thus assumes that "the stability of a ring of electrons rotating round a nucleus is secured through the above condition of the universal constancy of the angular momentum, together with the further condition that the configuration of the particles is the one by the formation of which the greatest amount of energy is emitted".
For systems of a more complicated constitution as having plural nuclei, Bohr makes use of the following theorem. Namely, in every system consisting of electrons moving in circular orbits with a velocity small compared with the velocity of light, the kinetic energy will be equal to half the absolute value of the potential energy. From this theorem, it follows that the total energy emitted by the formation of the system is equal to the kinetic energy of the electrons in the final configuration. In this way, Bohr is led to assume, "in analogy with the case of a single ring", that "corresponding to any configuration of equilibrium a series of geometrically similar, stationary configurations of the system will exist in which the kinetic energy of every electron is equal to the frequency of revolution multiplied by $\tau h/2$".

Bohr considers that, in such stationary configurations, the one corresponding to the greatest emitted energy will have $\tau = 1$ for every electron. Further, because the relation $T/\omega = \pi M$ holds for an electron in a circular orbit, he makes a "simple generalization", assuming that in any molecular system consisting of nuclei and electrons, in which the nuclei are at rest and the electrons move in circular orbits, the angular momentum of every electron round the center of its orbit will in the permanent state of the system be equal to $h/2\pi$.

§5. The structure of atoms and molecules

Making use of the assumption mentioned above of the constancy of angular momentum of every electron in the permanent state, Bohr investigates the structure of atoms in Part II "Systems Containing only a Single Nucleus". In doing so, he assumes further the condition of stability that the total energy of the system in the permanent state is less than that in any state, in which every electron is in a configuration slightly different from the one in the permanent state and rotates with an angular momentum $h/2\pi$. He uses an expression for the total energy of the system similar to that in the case of Nagaoka or Nicholson. It is used in a different way, however, to study the stability of the system on the basis of the quantum nature of angular momentum.

In order to see the stability of a ring, $n$ electrons rotating round a nucleus of charge $N e$ are supposed to be slightly displaced, in the direction perpendicular to the plane of the ring. Every electron in the displaced configuration is assumed to rotate with the angular momentum $h/2\pi$ in a circular orbit parallel to the original orbit. While the kinetic energy of the system is unchanged by

---

65) N. Bohr, Ref. 49) on p. 82 in the present chapter.
the displacements, the potential energy of the system is changed. From the requirement for this change to be positive, it results that

\[ N > p_{n,0} - p_{n,m}, \]  

(84)

where \( p_{n,k} \) is given by

\[ p_{n,k} = \frac{1}{8} \sum_{s=1}^{n-1} \cos \left( \frac{2k}{n} \frac{\pi}{s} \right) \csc^2 \left( \frac{s\pi}{n} \right), \]

and \( m \) is the value of \( k \) that makes \( p_{n,k} \) minimum.

Making use of Eq. (84), Bohr concludes that in the case of \( N = n \) any stable ring cannot exist unless \( n < 8 \), and that the number of the electrons which can rotate stably in a single ring is \( n = 10 \) for \( N = 20 \), \( n = 13 \) for \( N = 13 \) and \( n = 15 \) for \( N = 60 \).

In order to treat the case of electrons rotating in a number of rings, Bohr considers a uniform charge distribution along a ring. Supposing that \( n_1, n_2, \ldots \) electrons rotate in concentric rings of radii \( a_1, a_2, \ldots \) respectively in a common plane, he calculates the centripetal force exerted on the electrons in one ring by the nucleus and the electrons in the other rings, in the approximation of uniform charge distribution in the latter rings. From the balance of the centrifugal and centripetal forces, \( a_1, a_2, \ldots \) can be obtained by an iterative approximation. For such displacements of the electrons perpendicular to the plane of the rings as to produce no displacement of the centers of mass of the electrons in every ring, the condition of stability is given by Eq. (84). But, \( N \) is replaced with a value corrected for the effect of the other rings. If the electrons in a ring are displaced in an identical direction perpendicular to the plane of the ring, the electrons in the other rings make compensatory displacements. In this case, the condition of stability cannot be solved analytically.

As a result of a numerical study, Bohr finds that the number of electrons which the innermost ring can contain is approximately equal to the number of electrons in the case of a single ring. On the other hand, the number of electrons which can be contained in an outer ring is found to be considerably smaller than this value that is obtained from Eq. (84) with \( N \) replaced by the number of total charges inside the ring under consideration.

Bohr denotes the configuration in which \( n_1, n_2, \ldots \) electrons rotate round the nucleus of charge \( Ne \) in successive rings by \( N(n_1, n_2, \ldots) \), and the energy
of the system in this configuration by $W[N(n_1, n_2, \ldots)]$. For hydrogen atom, $W[1(1)]$ is obtained to be 13 volt. He denotes this by $W_0$.

In the case of the helium atom, if the configuration $2(1, 1)$ is supposed to be the permanent state, the two electrons may come so near to each other that deviations from circular orbits would be very great, because they rotate with identical angular momentum. Bohr therefore adopts the configuration $2(2)$ for the permanent state. Since $W[2(2)] - W[2(1)] = 2.13W_0$, he concludes that both the electrons in a helium atom are more firmly bound than the electron in a hydrogen atom. Nevertheless to say, for the correct understanding of the configuration $2(2)$ one should await the discovery of the electron spin in 1925.

For the lithium atom, Bohr gets $W[3(3)] = 17.61W_0$ and $W[3(2, 1)] = 16.02W_0$. He argues, however, that the configuration $3(2, 1)$ should be the permanent state by reason of the chemical properties of this element. Because in the configuration $3(2, 1)$ the ratio between the radii of the second and first rings is found to be larger than in the case of the helium atom, and two electrons are arranged symmetrically in the first ring, Bohr expects the effect mentioned for the configuration $2(1, 1)$ of the helium atom will be weak. He thinks also that the result $W[3(2, 1)] - W[3(2)] = 0.89W_0$ shows that the outer electron in the configuration $3(2, 1)$ is bound even more lightly than the electron in the hydrogen atom, and this may explain the apparent tendency of the lithium atom to become a positive ion.

For the case $N \geq 4$, Bohr finds $W[N(4)] < W[N(2, 2)]$. Since the ratio between the radii of the rings in the configuration $N(2, 2)$ approaches unity when $N$ becomes large, he supposes that the energy emitted by an eventual confluence of the rings will be great. It cannot be determined by his theory only, in which element four electrons are arranged in the innermost ring for the first time. Bohr looks for the answer to this in the chemical properties of elements. He observes also that $W[N(4, 4)] < W[N(8)]$ for $N < 14$, and further that the configuration $N[8]$ is stable for displacements perpendicular to the plane of the ring when $N \geq 10$.

After making analyses of these kinds, Bohr is led to the following possible scheme for the arrangement of electrons in light atoms.

He says that in this scheme the number of the electrons in every outermost ring is chosen so as to fit the ordinary valence of the respective element.

About the production of characteristic X-rays, Bohr considers the case in which an electron in the innermost ring is removed by impact of cathode-ray particles, for example, and the characteristic X-ray is emitted during the
settling down of the system. The minimum velocity \( u \) necessary for a cathode-ray electron to produce the characteristic X-ray is equal to the velocity of the electron in the ring. Putting \( E = Ne \) and \( \tau = 1 \), one gets from Eq. (72),

\[
u = \left(\frac{2\pi e^2}{h}\right) N = 2.1 \times 10^8 N \text{ cm/s},
\]

inversely to the case in Bohr's paper which we have quoted in § 1 of the present section. If one puts \( N = A/2 \), \( A \) being the atomic weight, this is in accordance with the experimental result \( A \times 10^8 \text{ cm/s} \) due to Whiddington\(^{66}\) for the elements from Al to Se.

In Part III "Systems containing Several Nuclei",\(^{67}\) Bohr discusses the constitution of molecules. For a system consisting of two nuclei of charge \( Ne \) and \( n \) electrons, he supposes that the electrons rotate in a ring round the axis connecting the nuclei. In the permanent state, the angular momentum of every electron is assumed to be \( h/2\pi \). From the balance between the centrifugal and centripetal forces acting on the electrons, he finds that the system is in equilibrium when the distance between the nuclei \( b \) is \( \left\{\frac{(4n/N)^{2/3} - 1}{2}\right\}^{-1/2} \) times the radius of the ring \( a \). From the stability of the system for displacements of the electrons perpendicular to the plane of the ring, he concludes that the system is stable only if \( N = 1 \) and \( n = 2 \) or \( 3 \), and this case corresponds to the hydrogen molecule.


\(^{67}\)N. Bohr, Ref. 50) on p. 82 in the present chapter.
The Bohr Model of the Atom

Denoting the radius of the orbit of the electron in the hydrogen atom by $a_0$, Bohr gets for the radius of that in a hydrogen molecule $a = 0.95a_0$, for the distance between the two nuclei $b = a/\sqrt{3}$ and for the energy of the hydrogen molecule $W = 2.20W_0$. $W - 2W_0$ amounts to $6.0 \times 10^4$ cal per mol. He says that this is "considerably less than" the value $13 \times 10^4$ cal per mol observed by Langmuir\(^{68}\) for the heat produced by the binding of two hydrogen atoms into a hydrogen molecule. In the case of large $N$, Bohr supposes that almost all electrons rotate separately round the two respective nuclei, and 2 or 3 outer electrons rotate round the axis connecting each pair of two nuclei.

In this way, Bohr developed the concept of the quantum nature of angular momentum in connection with the constitution of the hydrogen atom, derived spectral formulae obeying the combination rule from the quantum nature of angular momentum, and studied the constitution of atoms and molecules. In both Nicholson's and Bohr's work, however, the angular momentum considered was that of an electron rotating in a ring, the plane of which was taken to be the standard plane, and was a component of the angular momentum as a vector in the space. To see the quantization of the angular momentum as a vector, we should wait for Sommerfeld's quantum condition, which will be discussed in the next chapter.

§6. Moseley's verification of the Bohr model

Although the above scheme of configuration of atomic electrons given by Bohr was later greatly corrected, Moseley investigated in Part I\(^ {69}\) and Part II\(^ {70}\) of his paper "The High-Frequency Spectra of the Elements" the relation between the atomic number and frequency of characteristic X-rays $K_\alpha$ and $L_\alpha$, for many elements from Al to Au, and confirmed the fundamental structure of atom in the Bohr model that atoms should have stationary states due to the quantum nature of angular momentum.

In this paper, Moseley deduces the wavelengths of the characteristic X-rays by applying Bragg's formula to their reflection from a cleavage face of a crystal of potassium ferocyanide. Of the two K-lines of the characteristic X-ray, the stronger is called $K_\alpha$. About the frequency $\nu$ of $K_\alpha$, Moseley compares the

---


\(^{69}\)H. G. J. Moseley, Phil. Mag. 26 (1913), 1024.

\(^{70}\)H. G. J. Moseley, ibid. 27 (1914), 703.
The experimental value defined by

\[ Q = \sqrt{\nu / \frac{4}{3} \nu_0} \]

with the value expected from the Bohr model. \( \nu_0 \) is the fundamental frequency that is determined from the Rydberg constant and is given by \( \nu_0 / c = 109,720 \), \( c \) being the light velocity. He finds that \( Q \) increases by a constant amount as we pass from one element to the next, in accordance with the chemical order of the elements in the periodic table. Except for the cases of nickel and cobalt, this is also the order of the atomic weights. Moseley says that “we have here a proof that there is in the atom a fundamental quantity, which increases by regular steps as we pass from one element to the next”, and “this quantity can only be the charge on the central positive nucleus, of the existence of which we already have definite proof”. He says also that “we can confidently predict that in the few cases in which the order of the atomic weight \( A \) clashes with the chemical order of the periodic system, the chemical properties are governed by \( N \)”. \( N \) is “the number of the place occupied by the element in the periodic system”, namely the atomic number. Moseley thinks that “the very close similarity between the X-ray spectra of the different elements shows that these radiations originate inside the atom, and have no direct connection with the complicated-light spectra and chemical properties which are governed by the structure of its surface”.

The experimental value of \( Q \) can be expressed as

\[ Q = N - k, \]

where \( k \) is a constant, and “if \( N \) for calcium is really 20 then \( k = 1 \)”. Since there is good reason to believe that the K-lines come from the innermost ring of electrons, one has

\[ m \omega^2 r = \left( \frac{e^2}{r^2} \right) (N - \sigma_n), \]

where \( m \) is the mass of an electron, \( \omega \) the angular frequency of the electrons in the innermost ring, \( r \) the radius of their orbit, and \( \sigma_n \) is a small term identical with Bohr’s \( s_n \). Moseley’s equation is therefore equivalent to Eq. (80). Comparing the elements of atomic numbers \( N \) and \( N + 1 \), one gets

\[ (\omega^2 r^2)_{N+1} - (\omega^2 r^2)_N = \text{const}. \]
Because one has, on the other hand, from the experimental result of $Q$

$$\nu_{N+1}^{1/2} - \nu_N^{1/2} = \text{const},$$

one gets

$$\omega^2 r^3 / \nu^{1/2} = \text{const}.$$  

Since "for the types of radiation considered by Bohr... $\nu$ is proportional to $\omega$", it results that $\omega^3/2 r^3$ and therefore $m \omega r^2$, the angular momentum of an electron, is the same for all the different atoms. Moseley states that "thus we have an experimental verification of the principle of constancy of angular momentum, which was first used by Nicholson and is the basis of Bohr's theory of the atom".\(^{71}\)

He discusses that "it is evident that $k = \sigma_n$", and "if then $k = 1$, it is suggested that the ring contains 4 electrons, for $\sigma_4 = 0.96"$. This suggestion is not the same as the finding in the later development in the study of

\(^{71}\)About the nature of angular momentum of being an integer multiple of $h/2\pi$, A. Sommerfeld mentions in his paper [Ann. d. Phys. 51 (1916), 1], which we shall treat in the next chapter, that "it builds after the precedent by Nicholson the foundation of Bohr's theory of spectral lines." Also, A. Rubinowicz writes in Kap. 1, Bd. 24 of "Handbuch der Physik" (Springer, 1933), "Before Bohr this condition was already used by Nicholson in an attempt to explain the spectra of nebulas and solar coronas."

M. Jammer states in his book cited above, Ref. 54), on p. 88 in the present chapter, that "it should also be noted that Nicholson's anticipations of some of Bohr's conclusions were based, as Rosenfeld pointed out, on a most questionable and often even fallacious reasoning". Jammer says only that Nicholson worked "on the following assumption, that the ratio between the energy of the system and the frequency of rotation of the ring of electrons is given by an integral multiple of Planck's constant", but does not note that Nicholson pointed out and attached importance to the equivalence of this ratio to the angular momentum of electrons in a ring.

This is an incomprehensible opinion, that is not only in disregard of the descriptions of Moseley, Sommerfeld and Rubinowicz, but also is giving no eye to the words of Bohr himself cited above in § 3. The expression in Jammer's words quoted above "based on a most questionable and often even fallacious reasoning" is also very strange from the viewpoint of the history of science. Nicholson was not a student who answered questions about the Bohr model of the atom. Jammer stresses the importance of the Bohr model too much, to lose sight of the significance that Nicholson had arrived at the concept of the quantum nature of angular momentum before Bohr. We should not only grasp the definite difference between the atomic models of Nicholson and of Bohr, but also understand a great role that the quantum nature of angular momentum found by Nicholson played in the Bohr model.

Jammer makes a grave error in his discussion of Bohr's correspondence principle, by giving too much emphasis to Bohr's work [cf. Sec. 3, §3, p. 157, and Footnote 52), Sec. 4, §2, p. 161 of the next chapter].
atomic structure. Since the numerical agreement between \( \nu_0 \) and the factor \( 2\pi^2 e^4 m / h^3 \) which appears in the Bohr model is very close, Moseley expresses his experimental result of the K\(_\alpha\)-lines in the form

\[
\nu = \left( \frac{1}{2^2} - \frac{1}{3^2} \right) \frac{2\pi^2 e^4 m}{h^3} (N - 1)^2.
\]

The experimental result of the L\(_\alpha\)-lines is similarly expressed as

\[
\nu = \left( \frac{1}{2^2} - \frac{1}{3^2} \right) \frac{2\pi^2 e^4 m}{h^3} (N - 7.4)^2.
\]

If the experimental values of \( \nu^{1/2} \) are plotted graphically as a function of \( N \), they should be on a straight line. Moseley shows that this is actually the case for the known elements from Al to Au, on the assumption of \( N = 13 \) for Al. The order of the atomic numbers is found to be the same as that of the atomic weights, except where the latter disagrees with the order of the chemical properties. He states that “three, and only three, more elements are likely to exist between Al and Au”, and “as the X-ray spectra of these elements can be confidently predicted, they should not be difficult to find”.

It is well known that this diagram later became known as Moseley’s diagram and played an important role in atomic physics.

The verification of the existence of stationary states in the Bohr model was obtained through experiments on ionization energies of atoms. Franck and Hertz\(^ {72} \) measured in 1914 the acceleration voltage of electron-rays necessary to ionize mercury vapor. They found that up to a critical voltage of 4.9 volts there occurred only elastic scattering, and “the energy of the electron-rays of 4.9 volts was quite equal to the energy quantum of the resonant light of mercury of wave length 253.6 \( \mu \mu \)”. They considered this value of 4.9 volts to be the ionization voltage of mercury. Bohr\(^ {73} \) pointed out, however, that the energy of 4.9 volts “must correspond to a transition from the normal state of the atom to another stational state” supposed in the Bohr model.


\(^{73}\) N. Bohr, Phil. Mag. 30 (1915), 394.
Chapter 3

The Quantum Condition, Transition Probability and Correspondence Principle
Rutherford-Geiger (1908)
The statistical law of α-decay

Sommerfeld (1911)
For an elementary process
\( \int L \, dt = h/2\pi. \)

Sommerfeld (1915)
For rotational motion
\( \int pdq = nh. \)

Wilson (1915)
For each degree of freedom,
\( \int pdq = nh. \)
No treatment of angular momentum.

Ishiwara (1915)
The average of \( \int pdq \) per degree of freedom = \( nh \).
Consideration of angular momentum, to get a formula like Balmer's.

Debye (1914)
For a general oscillator
\( \iint dp \, dq = h. \)

Ehrenfest (1913)
For a rotating body
\( \iint dp \, dq = h. \)

Planck (1911)
For an oscillator
\( \iint dp \, dq = h. \)

Planck (1915)
For a system of \( f \) degrees of freedom
\( \int \cdots \int dp, \cdots dq = (nh)^f. \)
Debye (1916) Application of Sommerfeld's quantum condition to Zeeman effect.

Epstein (1916) Application of Sommerfeld's quantum condition to Stark effect.


Quantization of direction. Calculation of Zeeman effect.

Ehrenfest (1916) Adiabatic hypothesis.

Rubinowicz (1918) Derivation of the selection rule from conservation of angular momentum between electron and radiation.

Bohr (1918) The correspondence principle and selection rule, from the supposition that intensities of radiation are given by Fourier components.

Schwarzschild (1916) Application of Planck's quantum condition to Stark effect.


Notes: The numbers above or under a frame indicate Sec. and Subsec. 2 - means Chap. 2.
Chapter 3

The Quantum Condition, Transition Probability and Correspondence Principle

Theme in Chapter 3

Sommerfeld derived the Balmer formula of spectral lines by making use of the quantum condition. It was, however, the result of a process in which, by grasping the quantum nature of angular momentum as the quantum condition, he applied it to the rotational motion of an electron in an elliptic orbit, but arrived only at an energy of the electron dependent continuously on the eccentricity of the orbit, so that in order to quantize the eccentricity he "translated verbally" the quantum condition from the azimuthal coordinate to the radial coordinate. This process has not been told in almost all the books on the history of quantum mechanics.

To explain the fine structure of spectral lines, Sommerfeld treated the relativistic Kepler motion. Because for physically permissible orbits there was a minimum value $p_0$ of the angular momentum, he used the quantum condition in two ways, but could not help throwing away the "inevitably natural" way, because of the disagreement of its result with experiments. This point shows the limitation of the quantum condition. It has almost been forgotten that he introduced the fine structure constant $\alpha$ as the ratio of $p_0$ to $\hbar/2\pi$.

Spectral lines predicted from the energies of atomic stationary states with the use of the combination rule were too many and various that Sommerfeld was discouraged, saying "Does one quote only such cases as the confirmation of the combination principle, where observed combination lines agree well with calculated ones?" He limited the number of combination lines by putting a "provisional rule", but did not get satisfactory results.

It made a crucial turning point that Einstein introduced the transition probability into the theory of radiation. He supported, in his theory, emphatically the view that momentum conservation should be effective in the interaction between an atom and radiation. This viewpoint played a significant role in the later development of the quantum theory.
Bohr generalized his idea shown in his model of the atom and presented the correspondence principle, which assumed that, for transitions between stationary states with large quantum numbers, the transition probability would be determined by the Fourier components of the electric dipole moment of the system. It is well known that the correspondence principle gave rise to the selection rule. It should not be overlooked, however, that the calculated intensity of a spectral line was remarkably different whether the Fourier components were taken of the initial state or of the final one, showing the limitation of the correspondence principle in this meaning.

1. The Quantum Condition and Atomic Model

§1. Wilson’s quantum condition

The Bohr model of the atom was experimentally confirmed, as we have seen in the last chapter, by Moseley’s analysis of characteristic X-rays and Franck–Hertz’s measurement of ionization voltages, so that it became clear that electrons in an atom stay in discrete stationary states. There arose consequently the problem of how the stationary states of atomic electrons were determined. In the Bohr model of the hydrogen atom, it was made clear that the angular momentum of an electron describing a circular orbit should be an integer multiple of \( h/2\pi \), on the basis of the suggestion of Nicholson of the quantum nature of angular momentum.

For more general stationary orbits, what condition would determine these discreteness? Efforts to investigate such a condition, namely the one that later became known as the quantum condition, were made one after another by W. Wilson, J. Ishiwara and A. Sommerfeld in 1915.

In his paper “The Quantum Theory of Radiation and Line Spectra”,\(^1\) Wilson attempts at “deducing the results of Planck and Bohr from a single form of quantum-theory”. He formulates his theory on the following hypothesis.

Firstly, he supposes that “interchanges of energy between dynamical systems and the aether, or between one dynamical system and another, are ‘catastrophic’, or discontinuous in character”. This is to assume that “each system behaves as a conservative one during certain intervals, and between these are relatively short interval, during which definite amounts of energy may be emitted or observed”.

\(^1\)W. Wilson, Phil. Mag. 29 (1915), 795.
Secondly, he supposes that "the motion of a system in the intervals between such discontinuous energy exchanges is determined by Hamilton dynamics as applied to conservative systems". He defines the steady state by speaking of "a system, during such an interval, as being in one of its steady states".

These assumptions of Wilson correspond to the fundamental assumptions in the Bohr model, but Wilson adopts the Hamilton dynamics for "the ordinary mechanics" mentioned by Bohr in his fundamental assumptions. That is, let \( q_1, q_2, \ldots; p_1, p_2, \ldots \) be Hamilton's position and momentum coordinates of a system in a steady state, and \( L \) the kinetic energy. \( L \) is assumed to be given by a homogeneous quadratic form of \( q_1, q_2, \ldots \). If \( L \) contains terms such as \( q_r q_s (r \neq s) \), they are supposed to be eliminated by a linear transformation.

Thus, one has

\[
L = \frac{1}{2} A_1 \dot{q}_1^2 + \frac{1}{2} A_2 \dot{q}_2^2 + \cdots + \frac{1}{2} A_n \dot{q}_n^2. \tag{1}
\]

Further, because

\[
2L = \dot{q}_1 \frac{\partial L}{\partial q_1} + \dot{q}_2 \frac{\partial L}{\partial q_2} + \cdots + \dot{q}_n \frac{\partial L}{\partial q_n}, \tag{2}
\]

putting \( L_1 = A_1 \dot{q}_1^2 / 2, L_2 = A_2 \dot{q}_2^2 / 2, \ldots, L_n = A_n \dot{q}_n^2 / 2, \) one gets

\[
2L_1 = \dot{q}_1 p_1, \quad 2L_2 = \dot{q}_2 p_2, \ldots, \quad 2L_n = \dot{q}_n p_n. \tag{3}
\]

When the system is in a steady state, if it is assumed that the system has "a period \( 1/\nu_1 \) corresponding to \( \dot{q}_1, 1/\nu_2 \) corresponding to \( \dot{q}_2, \) and so on", there results that

\[
2 \int L_1 dt = \int p_1 dq_1, \quad 2 \int L_2 dt = \int p_2 dq_2, \quad \text{etc.} \tag{4}
\]

Then, Wilson supposes, as the third hypothesis, that "the discontinuous energy exchanges always occur in such a way that the steady motions satisfy the equations"

\[
\int p_1 dq_1 = \rho h, \quad \int p_2 dq_2 = \sigma h, \quad \int p_3 dq_3 = \tau h, \ldots. \tag{5}
\]

In these equations, \( \rho, \sigma, \tau, \ldots \) are integers, \( h \) is the Planck constant, and the integrations are over the respective periods.

Wilson's third hypothesis reminds us of Sommerfeld's hypothesis of action integral, which we have treated in the last chapter. Wilson's hypothesis would
correspond to an application of Sommerfeld's hypothesis to the steady motion until the occurrence of a "catastrophic" energy exchange, in which one period of the steady motion is put equal to the time of action. The reason why Wilson chooses the relation between the action integral and \( h \) as Eq. (5) is to make it possible, as is seen just below, to obtain the same result for a resonator as Planck's.

Now, putting \( q = R \cos(2\pi vt - \theta) \) for Planck's oscillator, one gets its energy \( E \) as

\[
E = 2\pi^2 \nu^2 m R^2,
\]

where \( m \) is the mass of the oscillator. Because

\[
\int_t^{t+1/\nu} p dq = 2\pi^2 \nu m R^2,
\]

one has from Eq. (5)

\[
E = \rho \nu.
\]

In the case of the Bohr model of the hydrogen atom, because \( L \) is constant in a steady state, one has from Eq. (4)

\[
2L \int_t^{t+1/\nu} dt = \rho \hbar,
\]

that is

\[
L = \rho \nu/2. \tag{6}
\]

Since in this case \( L \) is identical with Bohr's \( W \), Wilson says that Eq. (6) "expresses Bohr's principal hypothesis". This should be taken to indicate the correspondence of Eq. (6) with Eq. (71) in Sec. 4, §2 of the last chapter. About the energy emitted by an atom "in passing" from one steady state to another, Wilson mentions that "it must be equal to \( r_1 \nu_1 + r_2 \nu_2 + \cdots \), where the \( r \)'s are integers, not necessarily all positive, and \( \nu_1, \nu_2, \ldots \) are the frequencies of the corresponding aether vibrations". According to him, his theory "therefore includes this second assumption of Bohr as a special case".

It may be that Wilson overlooks the inconsistency, mentioned in Sec. 4, §2 of the last chapter, between Bohr's assumption \( W = \tau \hbar \omega/2 \) [Eq. (71), Chap. 2] and Bohr's another assumption [Eq. (75), Chap. 2], on account of Bohr's complex way of describing the logical development in his paper on the
atomic model. As we have already mentioned, following directly the thinking of Eq. (6) given above, namely Eq. (71) in the last chapter, it is impossible to arrive at Bohr’s frequency relation \( W_{T_2} - W_{T_1} = h\nu \). For this reason, Bohr has reinterpreted \( W = \tau h\omega /2 \) as \( M = \tau h/2\pi \) on the basis of the quantum nature of angular momentum introduced by Nicholson. If the importance of the angular momentum in the Bohr model were noticed, Wilson’s quantum condition Eq. (5) would give directly \( 2\pi M = \rho \) for the azimuthal angle \( \varphi \) and a constant angular momentum for the circular orbit.

§2. Ishiwara’s quantum condition and the spectral formula

Ishiwara, in his 1915 paper “The Universal Meaning of the Action Quantum”,\(^2\) takes notice also of the results of Planck, Sommerfeld and Bohr, and attempts to clarify the significance of \( h \) appearing in these theories. Referring to Planck’s hypothesis of quantum action, which we have treated in Sec. 1, §2 of the last chapter, Ishiwara says that “one is at present inclined to choose such a standpoint, that is to explain the appearance of \( h \) by the existence of definite finite domains of the state space”. He considers, however, that “the constant \( h \) is not only related to stationary radiative processes, but also is to be considered as well the outflow of a universal molecular character”. Then thinking that “in this relation Bohr’s atomic model is of great importance, in which the quantity, \( h/2\pi \) is ascribed a new meaning as the quantum of the angular momentum of the rotating electron”, he takes up the importance of the angular momentum in the Bohr model.

According to Ishiwara, even if Bohr’s hypothesis that the angular momentum of an electron in an atom should be an integer multiple of \( h/2\pi \) is in accordance with the fact, “this must not however be understood as the proper meaning of \( h \)”. This is “because the angular momentum is a vector quantity, whose absolute value in general cannot simply be composed linearly of its components”. It must be said that Ishiwara’s indication of this point is appropriate. For its solution, however, we should a wait Sommerfeld’s quantum condition, as we have mentioned at the end of Sec. 4, §6 in the last chapter.

Ishiwara continues to say that “from the same viewpoint, Sommerfeld’s proposition seems on the contrary to be of quite general validity”, attaching

---

\(^2\)J. Ishiwara, Publication of the Mathematical-Physical Society of Tokyo 8 (1915), 106 [received on Apr. 4, 1915].
importance to Sommerfeld's hypothesis on action integral.\(^3\) He thinks so because "it identifies \(\hbar\) with the very distinguished invariant quantity in mechanics, namely the action (apart from the factor \(2\pi\))."

From this standpoint, Ishiwara takes as his fundamental assumption the formulation that "the motions in nature always so take place that a splitting of each state plane \(p_k, q_k\) into those elementary domains of equal probability is allowed, of which the average value at a given point in the state space

\[
h = \frac{1}{j} \sum_{i=1}^{j} q_i p_i,
\]

is equal to a universal constant". In this equation, \(q_1, q_2, \ldots, q_j\) are the co-ordinates which determine the state of the system, and \(p_1, p_2, \ldots, p_j\) are the respective momenta corresponding to them.

Applying this fundamental assumption to an electron, he gets with the position vector \(r\) and the rest mass \(m_0\) of the electron

\[
h = \frac{1}{j} \int m_0 \dot{r} \cdot dr,
\]

where \(j\) is equal to 1, 2, or 3 as the motion is one-, two- or three-dimensional respectively. Every domain of equal probability is enclosed with two successive stationary motions. Consequently, if the integration in Eq. (8) is taken over the whole region surrounded by the stationary motion the right-hand side of Eq. (8) should be an integer times \(h\). Letting \(n\) be an integer, one gets

\[
njh = \int m_0 \dot{r} \cdot dr.
\]

With the use of the kinetic energy \(T = (1/2)m_0 \dot{r}^2\), this becomes

\[
njh = 2 \int_0^\theta T dt,
\]

where \(\theta\) is the period of the motion. Showing that for the central force proportional inversely to the distance squared Eq. (10) can be transformed to

\[
njh = 2 \int_0^\theta (T + V) dt,
\]

\(^3\)A. Sommerfeld, Phys. Zeit. 12 (1911), 1067. [In this paper the contents of Refs. 13) and 14) on p. 48 in Sec. 1, §3, Chap. 2 of the present volume are collected.]
where \( V \) denotes the potential energy. Ishiwara remarks that Sommerfeld's hypothesis of action integral \( \int_0^T (T - V) dt = h/2\pi \) "loses its ground". That is to say, Ishiwara thinks that the time integral of the Hamiltonian is connected with \( h \), instead of that of the Lagrangian.

In order to apply Eq. (9) to the Bohr model of the atom, Ishiwara considers an elliptical orbit with one focus at the nucleus, a major axis \( a \) and an eccentricity \( \varepsilon \). Letting the origin of the coordinate system be at the nucleus and the \((x, y)\) plane be the orbital plane, he has from Eq. (9) with \( j = 2 \)

\[
2nh = \int m_0 \dot{x}' dx' + \int m_0 \dot{y} dy ,
\]

(12)

where \( x' = x + \varepsilon a \). Carrying out the integrations with the use of the equation of the orbit and the area theorem, he derives

\[
h = \pi f / \sqrt{1 - \varepsilon^2} ,
\]

(13)

\( f \) being the angular momentum of the electron. Because for a circular orbit this becomes

\[
f = nh / \pi ,
\]

(14)

Ishiwara concludes that "Bohr's assumption is still not applicable to circular motion".

This difference comes from the fact that Ishiwara imposes a single quantum condition on the average of the action integrals for each degree of freedom as given by Eq. (7), instead of imposing quantum conditions separately on the action integrals for each degree of freedom as in Wilson's work. Since circular motion is substantially one dimensional, the relation \( f = nh/2\pi \) would be obtained in place of Eq. (14) by putting \( j = 1 \). Derivation of Bohr's result by imposing quantum conditions separately on the radial as well as rotational motions is done by Sommerfeld, as will be seen in the next subsection §3.

Now, Ishiwara then calculates the average kinetic energy \( \bar{T} \) of the electron. From the period of the elliptical motion given by \( 2\pi a^2 m_0 \sqrt{1 - \varepsilon^2} / f \), he obtains for the frequency \( \nu \), which is defined as the inverse of the period, the relation

\[
\frac{nh}{\pi} = 2\pi a^2 m_0 \nu .
\]

(15)

On the other hand, letting, \(-e\) and \(e'\) be the electric charges of the electron and the nucleus, respectively, there holds the relation
\[ a(1 - \varepsilon^2) = \frac{f^2}{m_0 e e'}. \] (16)

Eliminating \( f \) and \( a \) from Eqs. (13), (15) and (16), he gets

\[ \nu = \frac{\pi^2 m_0 e^2 e'^2}{2n^3 h^3}. \] (17)

Inserting this formula for \( \nu \) into the relation

\[ \tilde{T} = n h \nu, \]

which results from Eq. (10) with \( j = 2 \), Ishiwara derives

\[ \tilde{T} = \frac{\pi^2 m_0 e^2 e'^2}{2n^2 h^2}. \] (18)

For the frequency \( \nu_1 \) of the radiation emitted in the transition between two states, he thus obtains, using Bohr's frequency relation,

\[ \nu_1 = \frac{\pi^2 m_0 e^2 e'^2}{2h^3} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right). \] (19)

He mentions that for this formula to agree with the Balmer formula, "one has now to assume that the central charge of the hydrogen atom consists of two elementary electric charges".

Ishiwara's supposition of \( e' = 2e \) was not in agreement with the fact, and was due to the imposition of a single quantum condition on the average action integral. On the contrary, Wilson could not get a spectral formula of the Balmer type, though he applied the quantum condition on separate degrees of freedom. This was because Wilson did not notice the importance played by the quantum nature of angular momentum in the Bohr model, as Ishiwara did. It was in Sommerfeld's work that the quantum conditions for separate degrees of freedom were taken into account, together with the quantum nature of angular momentum, to arrive at the Balmer formula. It should be especially mentioned, however, that before Sommerfeld Ishiwara derived for the first time, by introducing his quantum condition, a spectral formula that is identical with the Balmer formula except for the factor of two mentioned above.
§3. The quantum theory of hydrogen atom due to Sommerfeld

In his 1915 paper "On the Theory of Balmer's Series", Sommerfeld took up the Bohr model of the hydrogen atom and imposed the quantum condition on the rotational and radial motions of the electron moving in an elliptical orbit, to derive the Balmer formula of spectral lines. In his next paper in 1916, "The Fine structure of Hydrogen and Hydrogen-like Lines", he treated the Kepler motion of the electron relativistically, to investigate the fine multiple structure of spectral lines, and to analyze the characteristic X-rays. These works, developed and improved in some points, were arranged into another paper in 1916, "On the Quantum Theory of Spectral Lines". It was in this paper that he found the quantization of direction by treating the motion of the electron three-dimensionally.

In the first paper cited above, Sommerfeld discusses the effect of the quantum condition on a periodic orbital motion. As has been mentioned in Sec. 1, §2 of the last chapter, Planck proposed the hypothesis of quantum of action that there should be in the phase space mutually independent elementary domains of equal probability, and presumed the condition

$$\int \int dqdp = h$$

(20)

for each elementary domain. Sommerfeld notes that from this condition the energy of an one-dimensional oscillator $W$ is given by $W = nh\nu$. He notes further that Planck's method was generalized in 1913 by Debye to an arbitrary one-dimensional periodic motion, and by Ehrenfest to a simple rotator.

In his lecture cited just above, Debye took up Grüneisen's rule that the volume expansion coefficient of a solid body is proportional to its specific heat at constant volume, and treated an aggregate of one-dimensional unharmonic oscillators. Debye emphasized that among the trajectories of various given energies in the $(q, p)$ plane, only such trajectories as to make the area between two neighbouring ones equal to $h$ should contribute to the calculation of the

---

6) A. Sommerfeld, Ann. d. Phys. 51 (1916), 1, 125.
partition function. He stressed that this could "characterize essentially the results of the quantum theory". As he pointed out, in the case of an unharmonic oscillator, the trajectories take forms different from elliptical in the case of Planck's resonator. In actual calculations, Debye used a kind of iterative approximation in which the unharmonic oscillation was treated as a harmonic one around a displaced equilibrium point, and applied Planck's method to obtain the average energy of such harmonic motion. Sommerfeld's attention is concerned in the principal side of Debye's discussions of the problem.

Ehrenfest treated the problem of the specific heat of gas of diatomic molecules. It was reported by Nernst at the 1st Solvay Conference in 1911 that the specific heat of gas of diatomic molecules at low temperatures took values smaller than those expected from the law of equipartition. Ehrenfest explained Eucken's observed values of hydrogen gas with the use of Planck's formula of the average energy of an oscillator, on the supposition that the rotational energy of a diatomic molecule should be an integer multiple of $h\nu/2$. In other words, he considered the case where the direction of a diatomic molecule carried out harmonic oscillation, on account of some external force. As the external force was weakened more and more, the motion of the molecule passes "adiabatically" into a uniform rotational motion. Denoting the rotational angle and angular momentum around the rotational axis of the molecule by $q$ and $p$ respectively, he suggested that the ellipse in the $(q, p)$ plane would pass into a pair of lines between $q = \pi$ and $q = -\pi$, whose values of $p$ were given by $p = \pm nh/2\pi$ from Planck's hypothesis of quantum of action for oscillators. This idea forms the basis of the adiabatic hypothesis, which will be treated in Sec. 4, §2 of the present chapter.

Now, Sommerfeld attempts to connect the quantum nature of angular momentum of the electron in the Bohr model with the hypothesis of quantum of action. For a particle of mass $m$ moving along a circle of radius $a$ with constant velocity, one has, taking its azimuthal angle $\varphi$ as the $q$-coordinate,

$$q = \varphi, \quad T = \left(\frac{m}{2}\right) a^2 \dot{\varphi}^2, \quad p = ma^2 \dot{\varphi} = \text{const}.$$  

A state of uniform rotation is represented in the $(q, p)$ plane by a line parallel to the $q$-axis. The ellipse in the case of an oscillator corresponds to a rectangle of base $2\pi$ in the present case. If the area of the rectangle is $h$, its height is $h/2\pi$. From this consideration, the angular momentum turns out to be

---

given by \( p = nh/2\pi \). Sommerfeld says that "the place of the discrete energy elements in the case of a mass point is therefore taken by Bohr's statement of the discrete angular momentum elements in the case of a rotating mass point".

In consideration of these examples, Sommerfeld tries to formulate generally the quantum hypothesis. He thinks of a group of infinite number of orbital curves in the \((q, p)\) plane, and considers the area between any two of them. If the curves are not closed, they can be closed by taking into consideration some periodicity or symmetry property, as in the case of a rotating mass point. Denoting the values of the \( p \)-coordinate on these curves successively by \( p_0, p_1, p_2, \ldots \), Sommerfeld puts

\[
\int \int dp dq = \int p_n dq - \int p_{n-1} dq = h, \tag{21}
\]

from the requirement that the area between the \((n-1)\)th and \(n\)th curves should be equal to \( h \). Assuming

\[
\int p_0 dq = 0
\]

for the first curve, he gets

\[
\int p_n dq = nh \tag{22}
\]

with the successive use of Eq. (21) starting from \( n = 1 \). Sommerfeld names this integral phase integral.

The momentum \( p \) corresponding to the coordinate \( q \) is given by \( p = \partial T/\partial \dot{q} \) in terms of the kinetic energy \( T \). If a system of coordinates is chosen by means of which \( T \) contains terms of the form \( \frac{1}{2} A_i \dot{q}_i^2 \) only, but not of the form \( \dot{q}_i \dot{q}_k \), one has

\[
\int pdq = \int A\dot{q} dq = \int Aq^2 dt > 0
\]

because \( A > 0 \). From this relation, \( n \) turns out to be a natural number since \( n = 0 \) is excluded. We see that Eq. (22) is the same as the one used by Wilson, and the phase integral is equivalent to the action integral.

Sommerfeld then calculates the kinetic energy of the Kepler motion, by making use of this quantum condition. Considering a hydrogen atom, he takes the polar coordinates \((r, \varphi)\) with the origin at the nucleus, whose mass is
assumed to be infinite. Integrating the equations of motion expressed in terms of the polar coordinates, he derives the kinetic energy $T$ and potential energy $V$:

$$T = \left(\frac{me^4}{2p^2}\right) (1 + \varepsilon^2 + 2\varepsilon \cos \varphi),$$

$$V = -\left(\frac{me^4}{p^2}\right) (1 + \varepsilon \cos \varphi),$$

where $\varepsilon$ is the eccentricity of the elliptic orbit and $p$ is the area constant, namely the angular momentum. $p$ is given by the semi-major axis $a$ and semi-minor axis $b$ of the ellipse and the average angular frequency $\omega$ as

$$p = mab\omega = ma^2\sqrt{1 - \varepsilon^2}\omega.$$  \hspace{1cm} (23)

The total energy of the electron becomes

$$W = T + V = -(me^4/2p^2)(1 - \varepsilon^2).$$ \hspace{1cm} (24)

Applying the quantum condition

$$\int pdq = p\int_0^{2\pi} d\varphi = 2\pi p = nh,$$ \hspace{1cm} (25)

Sommerfeld gets

$$W = W_n = -\frac{2\pi^2 me^4 1 - \varepsilon_n^2}{h^2} = -Rh\frac{1 - \varepsilon_n^2}{n^2},$$ \hspace{1cm} (26)

where $\varepsilon_n$ is the value of $\varepsilon$ determined from Eqs. (23) and (25), and $R$ is the Rydberg constant in Bohr's expression. From Bohr's second fundamental assumption, the frequency of light becomes

$$\nu = R \left(\frac{1 - \varepsilon_n^2}{n^2} - \frac{1 - \varepsilon_m^2}{m^2}\right).$$ \hspace{1cm} (27)

This however, gives, the Balmer formula only if circular orbits are considered, but gives rise, in general, to a band spectrum. Sommerfeld thinks therefore that "if we do not want to forbid the electron from describing also elliptic orbits besides circular ones the requirement turns out irrefutable that the eccentricity is to be calculated in a quantum way and to be connected with some integer value".
Thus, Sommerfeld "translates verbally" the quantum condition Eq. (25) from the azimuthal coordinate $q = \varphi$ to the radial coordinate $q = r$. The corresponding momentum is given by $p_r = \partial T / \partial r = m \dot{r}$, if the nucleus is assumed to be infinitely heavy. It follows that

$$\int p_r \, dr = \int m \dot{r} \, dr = \int_0^{2\pi} m \dot{r} \frac{dr}{d\varphi} \, d\varphi = n' h, \quad (28)$$

$n'$ being an integer. Carrying out the integration with the use of the equation of the ellipse $1/r = (me^2/p^2)(1 + \varepsilon \cos \varphi)$ and the relation $m \ddot{r} = (me^2\varepsilon/p) \sin \varphi$, which is obtained by integrating the equation of motion, Eq. (28) becomes

$$\int p_r \, dr = 2\pi p \left( \frac{1}{\sqrt{1 - \varepsilon^2}} - 1 \right) = n' h. \quad (29)$$

This, together with Eq. (25), gives

$$1 - \varepsilon^2 = \frac{n^2}{(n + n')^2}, \quad (30)$$

which brings forth, on account of Eq. (24),

$$W = -\frac{2\pi^2 me^4}{h^2} \frac{1}{(n + n')^2} = -\frac{R h}{(n + n')^2}. \quad (31)$$

Sommerfeld says that "this result is to the last degree surprising and is of a strikingly determinative character", and that "it seems to me impossible that such a precise and momentous result could be ascribed to an algebraic accident". From this expression of the energy, the Balmer series results at once. That is, considering, besides the orbit with the quantum numbers $n$ and $n'$ ("the final orbit of the electron"), a second one with the quantum numbers $m$ and $m'$ ("the starting orbit of the electron"), one gets

$$\nu = R \left\{ \frac{1}{(n + n')^2} - \frac{1}{(m + m')^2} \right\}, \quad (32)$$

which is nothing but the Balmer formula because $n + n'$ and $m + m'$ are integers. Sommerfeld points out that "the Balmer series seen in this new light depends on four integer numbers, but in the case of the hydrogen they reduce by chance, so to speak, to two integer numbers".

As we have mentioned in Sec. 1, §2 of the last chapter, Sommerfeld considered the importance of $h$ in connection with the action integral, and emphasized
the viewpoint that the existence of the molecule should be the function and result of the existence of the quantum of action. Being inspired by this suggestion, Nicholson investigated in 1912 the spectral lines of nebulae and solar coronae on the basis of the Nagaoka model of the atom, in which no assumption was made about atomic sizes, and concluded the quantum nature of angular momentum. Bohr, who verified in the same year the number of electron in the hydrogen atom and was able to know the substantialistic structure of the hydrogen atom, proposed in 1913 the Bohr model of the atom on the basis of the quantum nature of angular momentum, and assumed phenomenologically the existence of stationary states of electrons in an atom in order to explain the Balmer formula. Thus Sommerfeld gives, in this paper, launfulness based on the quantum condition due to the connection between the action integral and the quantum of action, to the existence of stationary states and spectral series. It is interesting to note that this process shows that the development of things takes place in a spiral way, so to say, not in a linear way.

Sommerfeld’s theory of the atomic structure and spectral lines opened the way to the system of theories that became known as the old quantum mechanics after the establishment of quantum mechanics. This system gave rise, however, to problems which should in turn be resolved.

2. The Quantum Condition and the Multiplicity of Spectral Lines

§1. The quantization of direction

Because Eq. (32) in the last section, which gives the Balmer formula of spectral lines, is determined by four quantum numbers, every spectral line with a given frequency possesses a multiplicity. This multiplicity comes from the fact that the energy of a stationary state determined by two quantum numbers \( n \) and \( n' \) depends on their sum as in Eq. (31). Putting \( N = n + n' \), Sommerfeld assumes that the number of pairs of \( n \) and \( n' \) allowed under a given \( N \) is equal to \( N \), excluding the case of \( n = 0 \).

In the case of \( n = 0 \), the angular momentum \( p \) of the electron is zero, and the elliptic orbit reduces to a linear orbit. This can be seen from the fact that the major and minor axes \( a \) and \( b \), respectively, of the elliptic orbit are given by

\[
a = \frac{h^2}{4\pi^2 me^2} (n + n')^2, \quad b = \frac{h^2}{4\pi^2 me^2} n(n + n')^2, \tag{33}
\]
which follow from the relations \( a = \frac{p^2}{me^2} / (1 - \varepsilon^2) \) and \( b = a\sqrt{1 - \varepsilon^2} \) and Eqs. (25) and (30). In the orbit with \( n = 0 \), the electron would run through the nucleus, and the velocity of the electron would become unlimitedly high at its approach into the nucleus, so that relativistic-theoretical corrections would be inevitable. According to Sommerfeld, "in any way we shall regard the orbit with \( n = 0 \) as highly problematic, and in the following take no account of it".

In this consideration, every spectral line arises in

\[
NM
\]

ways with \( M = m + m' \). Sommerfeld doubts, however, whether all these "transitions"\(^{10} \) are possible. Considering energy and energy quantum in the meaning of Bohr's frequency relation, Sommerfeld thinks that "energy is a positive quantity, and in a process which in itself goes by itself the energy can only decrease", and supposes from \( W_n - W_m > 0 \) that

\[
mm' > nn'.
\]

"At present treated are not energy quanta but action quanta", and linear or angular momenta take positive as well as negative signs and can increase as well as decrease. Since, however, in the phase integral \( p \) is connected with "an essentially positive quantity" through the multiplication of \( dq \), "the assumption itself suggests that the variation of these quantities in spontaneous transitions (freiwilligen Übergangen) takes place one-way as that of the energy, namely in decreasing sense". By this reasoning, Sommerfeld now assumes that "only such transitions are possible in which both the quantum numbers do not increase", and puts

\[
m \geq n, \quad m' \geq n'.
\]

"On this provisional assumption", the number of ways of occurrence of every spectral line is for \( M > N \) given by

\[
N(M - N + 1).
\]

In his 1916 paper"\(^{11} \) Sommerfeld mentions that these considerations "do not follow inevitably from the foregoing assumptions", and "are thereby hypothetical rather than guiding and depend on the respective experimental

\(^{10} \) Sommerfeld uses here the term "transition (Übergang)".

\(^{11} \) A. Sommerfeld, Ref. 6) on p. 114 in the present chapter.
circumstances". This shows that he is conscious of the phenomenological character of his theory of the multiplicity of spectral lines.

As one of such experiments, Sommerfeld takes the Stark effect. This is a phenomenon discovered by Stark,\textsuperscript{12}) in which each spectral line splits in an external electric field into several lines. Sommerfeld discusses in the first paper\textsuperscript{13}) in 1915 that "the electric field will influence various elliptic orbits in various ways and thereby originally coincident frequencies dissociate from each other", and that "this influence will be stronger in hydrogen than in other elements of which such a coincidence is not expected".

Sommerfeld compares Eqs. (34) and (34') with the number of split lines observed by Stark\textsuperscript{14}) for some spectral lines of hydrogen, as in the following table.

<table>
<thead>
<tr>
<th></th>
<th>p-Comp.</th>
<th>s-Comp.</th>
<th>Eq. (34')</th>
<th>Eq. (34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(_\alpha)</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>H(_\beta)</td>
<td>6 ~ 7</td>
<td>6 ~ 7</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>H(_\gamma)</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>H(_\delta)</td>
<td>7 ~ 8</td>
<td>7 ~ 8</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

In the table, p-Comp. denotes the number observed in the direction parallel to the electric field, and s-Comp. that observed in a direction orthogonal (\textit{senkrecht}) to the electric field. From the comparison, he concludes that a general parallelism between the calculated and observed numbers is not to be denied, especially if the standpoint of Eq. (34'), which is preferred from a general viewpoint, is adopted.

In order to make detailed comparisons it is necessary to calculate the orbits in an external electric field, but Sommerfeld says that he has not succeeded in it. In this case, the motion of the electron should be treated threedimensionally, because the direction of the electric field can be parallel as well as orthogonal to the plane of orbit. He discusses also the problem of generalizing the Balmer formula to elements other than hydrogen, and points out that the third coordinate \(z\) becomes necessary besides the \((r, \varphi)\) coordinates, because for such elements the electric field in an atom is less symmetric than for hydrogen.

\textsuperscript{12}) J. Stark, Berliner Beriche \textbf{1913}, 932.
\textsuperscript{13}) A. Sommerfeld, Ref. 4) on p. 114 in the present chapter.
\textsuperscript{14}) J. Stark, Göttinger Nachrft. \textbf{1914}, 427.
The three-dimensional calculation of the Stark effect is done, as is described in §3 of the present section, by Epstein and Schwartzschild independently in 1916, and three quantum numbers are introduced. Following their result, Sommerfeld treats the Kepler motion three-dimensionally in his 1916 paper, and arrives at the so-called quantization of direction by applying the quantum condition to the direction of the orbit, as is described in the following.

Sommerfeld points out that the direction of the orbital plane is physically indeterminate, and cannot be determined in the quantum theory also, because there is no standard direction. He selects therefore a standard of direction on the supposition that the motion of the electron in the electric field of the nucleus alone can be treated, by giving the space a specified direction by means of some imposed (internal or external) physical cause, and afterward reducing its strength to zero. Taking the direction to be that of the zenithal axis, he introduces spherical coordinates, \( (r, \theta, \varphi) \), and applies the quantum condition to each of them:

\[
\begin{align*}
\int p_\varphi d\varphi &= n_1 \hbar, \\
\int p_\theta d\theta &= n_2 \hbar, \\
\int p_r dr &= n' \hbar.
\end{align*}
\] (35)

The third phase integral is calculated in the same way as above, resulting in

\[
2\pi p \left( \frac{1}{\sqrt{1 - \varepsilon^2}} - 1 \right) = n' \hbar, \] (36)

where \( p \) is the area constant with respect to the orbital plane, but is here equal to the magnitude of the whole angular momentum. Letting \( \alpha \) be the inclination of the orbital plane from the equatorial plane, one has

\[
p_\varphi = p \cos \alpha,
\]

because the area theorem is applicable also with respect to the equatorial plane. From the first phase integral, one has, therefore,

\[
2\pi p \cos \alpha = n_1 \hbar.
\]

On the other hand, from the kinetic energy \( T = (m/2)(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2) \), \( p_\theta \) is given by \( p_\theta = \partial T / \partial \theta = m r^2 \dot{\theta} \), which, together with \( p = m r^2 \dot{\varphi} \), gives

\[
p_\theta = p \frac{d\theta}{d\varphi}, \quad \int p_\theta d\theta = p \int_0^{2\pi} \left( \frac{d\theta}{d\varphi} \right)^2 d\varphi.
\] (37)

15) A. Sommerfeld, Ref. 6) on p. 114 in the present chapter.
Integration of this with the use of the formula \( \cos \theta = \sin \alpha \sin \varphi \) results in
\[
\int p_{\varphi} d\varphi = 2\pi (1 - \cos \alpha) = n_2 h.
\] (38)

Eqs. (37) and (38) finally give
\[
2\pi \rho = (n_1 + n_2)h, \quad \cos \alpha = n_1/(n_1 + n_2).
\] (39)

Thus, one arrives at the quantization of direction. That is, with \( n = n_1 + n_2 \), \( \cos \alpha \) can take only \((n + 1)\) rational values. The total energy of the electron, however, depends only on \( n + n' = n_1 + n_2 + n' \), as is seen from Eqs. (25), (29), (31) and the first part of Eq. (39). Consequently, by the quantization of direction "no new line arises". But, Sommerfeld says that "the 'directional quantum number' introduced here gives the first possibility of relating the polarization to the spatial position of the orbital plane", and adds that in Epstein's paper, which we shall consider in §5 below, a rule for polarization in the Stark effect is obtained. Sommerfeld points out also that his calculations using spherical coordinates are not adequate for the Stark effect, but may be effective for the Zeeman effect in which the magnetic field concerned has an axial symmetry differently from the electric field concerned with the stark effect. The application of the quantization of direction to the Zeeman effect is made by Debye and by Sommerfeld in 1916. This will be discussed in §7 of the present section.

§2. The relativistic fine structure of spectral line

As we have seen in §1 above, in his 1915 paper Sommerfeld attempted to relate the multiplicity of the elliptic orbits of the electron with the dissociation of spectral lines in the Stark effect. He could not explain, however, the fine structure of the spectral lines in the case of no external field, because the energy in an elliptic orbit depends only on \( n + n' \), the sum of the azimuthal quantum number \( n \) and the radial quantum number \( n' \). It was known that many spectral lines were fine doublet or triplet lines, when observed with an apparatus of high resolution. For example, \( H_\alpha \) line in the Balmer series was known to consist of doublet lines with a difference in wavelength of 0.14 Å.

In his second paper,\(^\text{16}\) Sommerfeld treats the motion of the electron in a relativistic way, to calculate fine energy differences for various elliptic orbits.

\(^{16}\)A. Sommerfeld, Ref. 5) on p. 114 in the present chapter.
Because among the elliptic orbits with the same value of \( n + n' \), those of higher eccentricity have perihelions nearer the nucleus and thereby higher relativistic effect, the energy of the orbital motion is dependent on the eccentricity. In contrast to this viewpoint, Bohr\(^{17}\) calculated in his paper in 1915 relativistic corrections for circular orbital motions. Bohr attempted to explain the deviation of the "center of gravity" of the doublet lines from the Balmer formula, but the theoretical values accounted only for 1/3 of the deviations observed by Curtis\(^{18}\). Furthermore, because the intervals between the doublet lines were much larger than the deviations of the "center of gravity", Bohr remarked that "it seems hardly justifiable to compare the measurement of Mr. Curtis with any theoretical formula, unless the observed doubling of hydrogen lines is taken into account". He argued also that "possibly the lines were not true doublets, but the doubling observed was produced by the electric field in the discharge".

On Bohr's viewpoint, Sommerfeld comments that "we alter in principle the standpoint, taking up the following proposal: The doublet must have its ground on the finitely different discrete eccentricities of our 'quantized' ellipses". Let \( E \) denote the electric charge of the nucleus, and \(-e\) that of the electron. If the nucleus is at rest, the force acting on the electron is the Coulomb force. Since the area theorem is valid, one has, with the rest mass of the electron denoted by \( m_0 \),

\[
m r^2 \dot{\varphi} = \dot{p}, \quad m = m_0 / \sqrt{1 - \beta^2}, \quad \beta = u/c,
\]

from which one gets

\[
m \dot{\varphi} = \frac{p}{r^2}, \quad \frac{d}{dt} = \frac{p}{m r^2} \frac{d}{d\varphi}.
\]

Putting \( \sigma = 1/r \), one obtains from the equation of motion

\[
\frac{d^2 \sigma}{d\varphi^2} + \sigma = \frac{eEm}{p^2} = \frac{eEm_0}{p^2} \frac{1}{\sqrt{1 - \beta^2}}.
\]

From the time component of the equation of motion one has the law of energy conservation

\[
m_0 c^2 \left( \frac{1}{\sqrt{1 - \beta^2}} - 1 \right) - \frac{eE}{r} = W,
\]

\(^{17}\)A. Bohr, Phil. Mag. 19 (1915), 332.

$W$ being the total energy of the electron. On using this, Eq. (41) is rewritten as

$$\frac{d^2\sigma}{d\varphi^2} + \sigma \left[ 1 - \left( \frac{eE}{pc} \right)^2 \right] = \frac{eEm_0}{p^2} \left( 1 + \frac{W}{m_0c^2} \right).$$

Integrating this equation, one gets

$$\sigma = A \cos \gamma \varphi + B \sin \gamma \varphi + C,$$

where $A$ and $B$ are arbitrary constants, and $\gamma$ and $C$ are given by

$$\gamma^2 = 1 - \left( \frac{eE}{pc} \right)^2, \quad C = \frac{eEm_0}{\gamma^2 p^2} \left( 1 + \frac{W}{m_0c^2} \right). \quad (42)$$

The orbit is, therefore, an ellipse with rotating perihelion. The perihelion moves by an angle $2\pi(1/\gamma - 1)$ in the direction of the revolution for each revolution round the orbit. If the perihelion is initially at $\varphi = 0$, the orbit is represented by the equation $1/r = C(1 + \varepsilon \cos \gamma \varphi)$.

The magnitude of the movement of the perihelion, and consequently the degree of dissimilarity of the orbit from an ellipse, depends on the difference of $\gamma$ from 1. From Eq. (42) one has

$$\gamma^2 = 1 - \frac{p_0^2}{p^2}, \quad p_0 = eE/c, \quad (43)$$

which shows that $p_0$ is the limiting value of $p$. Since Sommerfeld adds some considerations of $p_0$ in his third paper,\(^{19}\) let us hereafter proceed by following this paper.

For $\varphi = 0$, because of the relations $\sigma = C(1 + \varepsilon)$ and $u = r\dot{\varphi}$, one obtains from the area theorem and the law of energy conservation

$$\frac{\beta}{\sqrt{1 - \beta^2}} = \frac{p\sigma}{m_0c} = \frac{eE(1 + \varepsilon)}{\gamma^2 pc} \left( 1 + \frac{W}{m_0c^2} \right), \quad (44)$$

$$\frac{1}{\sqrt{1 - \beta^2}} = \left[ 1 + \left( \frac{eE}{\gamma pc} \right)^2 \left( 1 + \varepsilon \right) \right] \left( 1 + \frac{W}{m_0c^2} \right). \quad (45)$$

Eliminating $\beta$ from these equations one gets

$$1 + \frac{W}{m_0c^2} = \left( \frac{p^2 - p_0^2}{p^2 - \varepsilon^2 p_0^2} \right)^{1/2}. \quad (46)$$

\(^{19}\)A. Sommerfeld, Ref. 6) on p. 114 in the present chapter.
Making use of this, the equation of the orbit can be written as

\[
\sigma = \frac{m_0 e E}{\sqrt{(p^2 - p_0^2)(p^2 - \varepsilon^2 p_0^2)}} \left[ 1 + \varepsilon \cos \left( \frac{\sqrt{p^2 - p_0^2}}{p_0} \varphi \right) \right].
\] (47)

When \( p = p_0 \) and \( \varepsilon \neq 1 \), one has \( \sigma = \infty \), or \( r = 0 \), so that the orbit reduces to the “zero orbit”. In this case, from Eqs. (44) and (45) one has \( \beta = 1 \), and the mass of the electron becomes infinite. Thus, \( p_0 \) makes “a natural lower limit” of the angular momentum. It can be seen that when \( p = p_0 \) and \( \varepsilon = 1 \) the orbit is either the zero orbit or a spiral circulating round the nucleus infinite times to approach it unlimitedly. Sommerfeld points out that in any way when the mass of the electron increases, the approximation of the nucleus being at rest becomes invalid. He mentions as follows — “I have believed that the manifoldness of lines in the spectrum of multi-line of hydrogen should be brought into a connection with the manifold form of the orbit in the limiting case described above. This thought failed because of the impossibility of bringing such orbits into agreement with the quantum conditions.” We shall treat this point in §3 below.

Now, the quantum condition for the phase integral with respect to \( q = \varphi \) is in the present case also given by

\[ 2\pi p = n\hbar. \] (48)

With respect to \( q = r \), one makes use of the relation

\[ p_r = m\dot{r} = m\dot{\varphi} \frac{dr}{d\varphi} = \frac{p}{r^2} \frac{dr}{d\varphi} = -p \frac{d\sigma}{d\varphi}, \]

which comes from the area theorem, to get

\[ n'\hbar = -p \int \frac{d\sigma}{d\varphi} dr = -p \int_0^{2\pi/\gamma} \frac{d\sigma}{d\varphi} \frac{dr}{d\varphi} d\varphi = p \int_0^{2\pi/\gamma} \left( \frac{1}{\sigma d\varphi} \right)^2 d\varphi, \]

by considering that \( \varphi \) changes from 0 to \( 2\pi/\gamma \) when the electron after passing through the perihelion comes again to it. Carrying out the integration by means of Eq. (47), one has for the radial quantum condition

\[ 2\pi \sqrt{p^2 - p_0^2} \left( \frac{1}{\sqrt{1 - \varepsilon^2}} - 1 \right) = n'\hbar. \] (49)
From Eqs. (46), (48) and (49) it follows that

\[ 1 + \frac{W}{mc^2} = \left\{ 1 + \frac{(\alpha E/e)^2}{[n' + n\sqrt{1 - (\alpha E/ne)^2}]^2} \right\}^{-1/2}. \]  

(50)

In this equation, \( \alpha \) is the ratio of \( p_0 \) for the hydrogen atom to the minimum angular momentum \( h/2\pi \) allowed by the quantum condition, which is given by

\[ \alpha = \frac{e^2}{c} \sqrt{\frac{h}{2\pi}} = 2\pi \frac{e^2}{hc} \sim 7 \times 10^{-2}. \]  

(51)

In Sommerfeld’s words, “the relativistic limit of angular momentum \( p_0 \) turns out to be very small in comparison with the quantum angular momentum \( p_n \)”, where \( p_n = nh/2\pi \). \( \alpha \) is nowadays called the fine structure constant. It is very interesting to see, in relation to the passage from Nicholson via Bohr to Sommerfeld, that \( \alpha \) has been introduced in connection with the angular momentum. It will be seen that \( \alpha \) plays an important role in the later development of quantum theory.

By making use of Bohr’s frequency relation, Sommerfeld derives

\[ \nu = \frac{m_0c^2}{h} \left\{ \left( 1 + \frac{\alpha^2}{(n' + \sqrt{n^2 - \alpha^2})^2} \right)^{-1/2} \right. \\
- \left. \left( 1 + \frac{\alpha^2}{(m' + \sqrt{m^2 - \alpha^2})^2} \right)^{-1/2} \right\}, \]

(52)

as the relativistic Balmer formula for the hydrogen atom. In this way, he has shown that “in view of the variable mass of the electron the energy becomes a little separated each for these \( n + n' \) various orbits”.

§3. The limitation of the quantum condition

In the above discussions, the quantum condition given by Eq. (22) in Sec. 1, §3 of the present chapter has been applied on the supposition of the smallness of \( p_0 \). If it is taken into account that the angular momentum cannot be smaller than \( p_0 \), the condition

\[ \int (p_n - p_0) dq = nh \]

(22')
The Quantum Condition, Transition Probability and Correspondence Principle

should be applied, instead of Eq. (22). If this condition is chosen, the azimuthal quantum condition becomes

\[ 2\pi(p_n - p_0) = nh, \quad p = p_0 + nh/2\pi. \]

As a result, one has for the hydrogen atom

\[ 1 + \frac{W}{m_0c^2} = \left\{ 1 + \frac{\alpha^2}{n' + n\sqrt{1 - 2\alpha/n}} \right\}^{-1/2}, \quad (50') \]

instead of (50).

If Eq. (50) is expanded in a power series in \( \alpha \), one has

\[ W = -\frac{Rh}{(n + n')^2} \left\{ 1 + \frac{\alpha^2}{n + n'} \left( \frac{1}{4} + \frac{n'}{n} \right) \right\}, \quad (53) \]

where \( R \) is the Rydberg constant. On the other hand, if Eq. (50') is expanded in the same way, one gets

\[ W = -\frac{Rh}{(n + n')^2} \left[ 1 - \frac{2\alpha}{n + n'} + \frac{\alpha^2}{n + n'} \left( \frac{9}{4} + \frac{n'}{n} \right) \right]. \quad (53') \]

Because Eq. (53') contains a term linear in \( \alpha \), it causes a large discrepancy from the Balmer formula, and thereby "it is denied even by the old and less exact measurements". Sommerfeld mentions that "this is surprising and regrettable, for it would be undeniably natural to let the quantum partition of the azimuthal \((p, q)\) plane start from the lower limit \( p = p_0 \) instead of from \( p = 0 \)."

Sommerfeld's indication of the theoretical contradiction contained in the quantum condition is very noticeable. Although the relativistic energy formula (53) is derived by applying the quantum condition to the relativistic orbit of the electron, the quantum condition given by Eq. (22) is not chosen because of logical inevitability. This shows that the quantum condition has limitations characteristic of a certain lawfulness in a substantialistic stage, and therefore the unification of the quantum theory and the relativistic theory given in the form of Eq. (53) has limitations of the same kind.

§4. The necessity of selection rules

Sommerfeld uses the relativistic energy \( W_{n,n'} \) obtained above as the spectral term "in the meaning of Ritz's combination principle" as in the Bohr model of the atom, to discuss the splitting of spectral lines. In doing so, he excludes
the term with $n = 0$ for "the geometrical reason" mentioned above. The spectral term with $N = n + n' = 2$ is a doublet, which has the difference in energy of

$$W_{1,1} - W_{2,0} = -(R\alpha^2/2^4)(E/e)^2,$$

($E$ being equal to $e$ for the hydrogen atom). Similarly, the spectral term with $N = 3$ "gives the cause for a triplet". A spectral line becomes in general a $NM$-multiplet on account of the combination of the terms of such kind ($M = m + m'$). It should be assumed, however, that "a part of these $MN$ components result" by imposing the condition that $m \geq n$ and $m' \geq n'$, as has been described in §1 of the present section.

Sommerfeld thinks that similar fine structures may be seen in spectral series of other elements which have hydrogen-like spectral terms. For example, because the $2p$ term of Li is given by $R/(2 - 0.041)^2$ and has as good an integer nature as the term $R/2^2$ of H, he infers "the similarity in the atomic field and in the respective orbits of the electron", and concludes that the $2p$ term of Li would be a doublet of a magnitude similar to that of H. He concludes also that the $3d$ terms of Li, Na and K are of triplet type because they are good in their integer nature. Sommerfeld "emphasizes that our discussion of the Li doublet builds a bridge to the explanation of the doublets of the alkalis, thereby of the D-doublets above all, and of other hydrogen-like terms".

He points out, however, that for the purpose of this explanation, special investigations must take decisive place in the atomic fields, by which new constants characterizing them would be introduced. When the integer nature of spectral terms is not good, the multiplicities of spectral lines have their origin in the deviation of the atomic field from the Coulomb field, and are of quite different magnitudes. Sommerfeld presents here an appropriate methodology by mentioning that "in their theoretical investigation, the task cannot be the prediction of the position and structure of the lines, but must consist in the determination of the nature of the atomic field, and thereby the structure of the atom, from the spectroscopic experiences". In order to pursue this subject, however, it turns out necessary to introduce the selection rule of spectral lines, as will be described below in Sec. 4.

Sommerfeld further discusses X-ray spectra, saying that "in the case of X-ray frequencies, especially those of heavy elements, multiplicities of macroscopic magnitude will appear", and "here the proof of our theory can take
place most certainly”. Kossel\textsuperscript{20} has made clear that the frequency difference between the X-rays $K_\alpha$ and $K_\beta$ is equal to the frequency of the X-ray $L_\alpha$. He has supposed that a characteristic X-ray is generated by the process, in which one of the electrons forming a group in an atom is removed from the atom, and another electron belonging to a different group makes a transition into the group to which the former electron belonged. On this supposition, he has considered that each absorption limit in the spectrum of absorbed X-rays results from the discontinuous change of the absorption coefficient of X-rays at the wavelength corresponding to the binding energy of an atomic electron.

Accordingly, for the X-ray spectrum Sommerfeld first takes the spectral terms in the same way as for visible light spectrum, and then derives X-ray frequencies on the basis of the combination rule. Denoting the spectral terms by $K, L, M, N, \ldots$, he represents $K_\alpha, K_\beta$ and $L_\alpha$ as $K - L, K - M$ and $L - M$, respectively.

The $K$ and $L$ terms are given according to Moseley\textsuperscript{21} by $R(Z - k)^2/1^2$ and $R(Z - l)^2/2^2$, respectively, in terms of the atomic number $Z$ and the Rydberg constant $R$, and “behaves exactly like hydrogen”. Sommerfeld concludes therefore that the $K$-term is singlet and the $L$-term is doublet, and there should be another $L$-term which is denoted by $L'$. The combinations $K - L'$ and $L' - N$ give rise to $K'_\alpha$ and $L'_\beta$ respectively. The frequency difference $\Delta \nu$ between $K_\alpha$ and $K'_\alpha$, as well as that between $L_\beta$ and $L_\alpha$, is given by $\Delta \nu/(Z - l)^4 = R \alpha^2/2^4$, namely the relativistic correction in terms of “the effective electric charge” $E = (Z - l)e$. The observational results due to Moseley and those Siegbahn\textsuperscript{22} show that this relation holds from $Z = 24$ to $Z = 92$.

According to Siegbahn, there are 13 X-ray lines which can be classified into the $L$-series. If they are supposed to be doublet lines corresponding to the combinations $L - X$ and $L' - X$ with $X = M, N, O$ and $P$, the pairs of lines do not satisfy the above relation. Sommerfeld introduces thus new terms $\Lambda$ and $\Lambda'$, and considers the combinations $\Lambda - X$ and $\Lambda' - X$ with $X = M$ and $N$. He supposes that the $\theta$-line in the $L$-series and the $\chi$-line in the $\Lambda$-series “make a very narrow doublet”.

To determine the terms $M, N, O$ and $P$, it turns out necessary to distinguish $M_K, M_L, N_K, N_L$, and so on, depending on whether they are derived from $K_\beta$ and $K_\gamma$ or $L_\alpha$ and $L_\beta, \ldots$. Because the application of the relativistic


\textsuperscript{21}H. Moseley, Refs. 69) and 70) on p. 99 in Sec. 4, §6, Chap. 2 of the present volume.

correction "to non-hydrogenlike terms would be valueless", Sommerfeld puts $M_K/R = (Z - m)^2/c^2$ and obtains $c = 2.78$ from its comparison with the observational results. The values of the constant $c$ for $M_L, N_K, N_L, O_L$ and $P_L$ are respectively $2.50, 3.09, 2.84, 3.25$ and $3.33$, rather bad in the integer nature. When the values of the square root of each term are plotted as a function of $Z$, the prolongation of the line representing $N_K$ passes between the lines representing $M_L$ and $N_L$, and that of the line representing $N_L$ between the lines representing $N_L$ and $O_L$. It turns out, therefore, that for large values of $Z$ "we cannot derive any of the L-lines from the K-lines by the combination rule".

From Siegbahn's observational results, the relation between the M-series and the L-series is similarly negative for heavy metals. In particular, the two strong lines $M_\alpha, N_\alpha$ cannot be obtained from any combination of L-lines. Sommerfeld concludes therefore that "the combination rule fails at the very place where it could most precisely be checked". Though he has treated spectral lines on the basis of the combination rule, by "regarding the quantization of electron orbits as experimentally and theoretically reliable", he now considers the combination rule not to be unconditional.

"Should, perhaps, the combination rule not have any absolute validity also in the visible region, in particular not so for non-hydrogenlike terms? Does one quote only such cases as the evidences of the combination rule, in which the observed combination lines are precisely in accordance with the calculated ones, while doesn’t one connect all the cases of discrepancy with the combination rule? In any way, as H. Konen treats in detail in his book ‘The Lighting of Gases and Vapours (Das Leuchten der Gase und Dämpfe)’, §67ff., of all the thinkable combinations only a small part is actually observed."

This difficulty pointed out by Sommerfeld showed that the quantum condition and the combination rule had not such launfulness as to be unconditionally connected with each other. The condition of connecting the quantum condition with the combination rule to develop the theory of spectral lines was the selection rule, which will be treated below in Sec. 4, §1 and §2 of the present chapter. With the use of the selection rule and Sommerfeld’s relativistic energy formula the analysis of X-ray spectra was brought into development. In the process of this development, however, the contradictory nature inherent in the relativistic energy formula which we have discussed above, became manifest as will be seen in Sec. 2, §3, Chap. 4 of the present volume.
§5. The quantum condition and the Stark effect

The work of applying the quantum condition to the calculation of the Stark effect was done in 1916 by Epstein and by Schwartzschild. Both of them noticed that the partial differential equation of Hamilton–Jacobi for an atomic electron in this case could be separated in some curvilinear coordinate systems.

In his paper "On the Theory of the Stark Effect", Epstein uses three quantum conditions of Sommerfeld's type, \( \int p_i dq_i = n_i \hbar \), "correspondingly to the three degrees of freedom of our electron". He gives the name "statistical orbit" to "the orbit that satisfies certain quantum conditions and is regarded as possible (or particularly probable) by itself from the quantum theory". If the electric charges of the electron and the nucleus are \( -e \) and \( +ke \) respectively, and the homogeneous external electric field is of strength \( E \) and is in the direction of the z-axis, the potential energy \( U \) in the system of the polar coordinates \((\rho, z, \varphi)\) is given by

\[
U = -\frac{\kappa e^2}{r} - eEz; \quad r = \sqrt{\rho^2 + z^2}.
\]

(54)

Epstein uses parabolic coordinates \((\xi, \eta, \varphi)\) defined by

\[
z = \frac{(\xi^2 - \eta^2)}{2}, \quad \rho = \xi \eta.
\]

(55)

In this coordinate system, \( U \) and the kinetic energy \( T \) are expressed as

\[
U = -\frac{4\kappa e^2 + eE(\xi^4 - \eta^4)}{2(\xi^2 + \eta^2)},
\]

(56)

\[
T = \frac{m}{8}(\xi^2 + \eta^2) \left( \xi^2 + \eta^2 + \frac{\xi^2 \eta^2}{\xi^2 + \eta^2} \varphi^2 \right).
\]

(57)

With the use of the momenta in this coordinate system \( p_\xi = \partial T/\partial \dot{\xi}, \ p_\eta = \partial T/\partial \dot{\eta} \) and \( p_\varphi = \partial T/\partial \dot{\varphi} \), \( T \) is rewritten to be

\[
T = \frac{m}{2m(\xi^2 + \eta^2)} \left[ p_\xi^2 + p_\eta^2 + \left( \frac{1}{\xi^2} + \frac{1}{\eta^2} \right) p_\varphi^2 \right].
\]

(58)

Because the Hamiltonian \( H = T + U \) does not contain the time \( t \) explicitly, he puts Hamilton's principal function \( S \) in the form \( S = At + W \) in terms of

---

Hamilton's characteristic function $W$, to get the partial differential equation of Hamilton–Jacobi

$$\frac{1}{2m(\xi^2 + \eta^2)} \left\{ \left( \frac{\partial W}{\partial \xi} \right)^2 + \left( \frac{\partial W}{\partial \eta} \right)^2 + \left( \frac{1}{\xi^2} + \frac{1}{\eta^2} \right)^2 \left( \frac{\partial W}{\partial \phi} \right)^2 \right\} - 4\kappa me^2 - meE(\xi^4 - \eta^4) \right\} + A = 0,$$

(59)
in which $-A$ has the meaning of the energy of the electron.

Epstein effects the separation of variables by putting

$$W = \sqrt{m\alpha} \phi + W_1(\xi) + W_2(\xi).$$

(60)

In this equation $\sqrt{m\alpha}$ represents the aerial constant of the motion around the $z$-axis. Putting the separation constant to be $me^2\beta$, he gets

$$\frac{dW_1}{d\xi} = \sqrt{m} \left\{ 2e^2(\kappa + \beta) - 2A\xi^2 + eE\xi^4 - \frac{\alpha^2}{\xi^2} \right\}^{1/2} \equiv \sqrt{m}f_1(\xi),$$

$$\frac{dW_2}{d\eta} = \sqrt{m} \left\{ 2e^2(\kappa - \beta) - 2An^2 - eEn^4 - \frac{\alpha^2}{n^2} \right\}^{1/2} \equiv \sqrt{m}f_2(\eta).$$

(61)

Noting the relations $p_\xi = \partial W/\partial \xi$, $p_\eta = \partial W/\partial \phi$ and $p_\phi = \partial W/\partial \phi$, Epstein takes as the three quantum conditions the following equations.

$$2\sqrt{m} \int_{\xi_1}^{\xi_2} \sqrt{f_1(\xi)}d\xi = n_1 h,$$

$$2\sqrt{m} \int_{\eta_1}^{\eta_2} \sqrt{f_2(\eta)}d\eta = n_2 h,$$

$$\sqrt{m} \int_0^{2\pi} \alpha d\phi = n_3 h.$$  

(62)

In the above equations, $\xi_1$ and $\xi_2$ are the limits of the $\xi$-coordinate which are determined by the zeros of $f_1(\xi)$, similar to the limits of the radial coordinate $r$ in the case of the Kepler motion treated by Sommerfeld. $\eta_1$ and $\eta_2$ are determined by the zeros of $f_2(\eta)$. Though the electron does not describe a closed orbit, the coordinates $\xi$ and $\eta$ have their respective lower and upper limits also in the present case. Epstein points out that "the problem treated belongs to a kind of motion that is called a libration by Charlier".\(^{24}\)

Supposing that the force exerted by the external field on the electron is small compared with the attractive force due to the nucleus, Epstein calculates

the effect in the first order of $E$. For this purpose, he introduces variables $u$ and $w$ defined by

$$u = 2A\xi^2/S_1 - C_1, \quad w = 2Ar^2/S_2 - C_2,$$

where

$$S_1 = \{e^4(\kappa + \beta)^2 - 2\alpha^2 A\}^{1/2}, \quad C_1 = e^2(\kappa + \beta)/S_1,$$

$$S_2 = \{e^4(\kappa - \beta)^2 - 2\alpha^2 A\}^{1/2}, \quad C_2 = e^2(\kappa - \beta)/S_2.$$

In the approximation to the first order of $E$, the limits of $u$ as well as those of $w$ are $\pm 1$. The third condition given in Eq. (62) is independent of $E$, and gives

$$\sqrt{m\alpha} = \frac{n^3h}{2\pi}.$$

For $A$ and $\beta$ in Eq. (64), Epstein writes

$$A = A_0 + \Delta A, \quad \beta = \beta_0 + \Delta \beta.$$

$A_0$ and $\beta_0$ are respectively the values of $A$ and $\beta$ for $E = 0$, while $\Delta A$ and $\Delta \beta$ represent respectively the first order effects on $A$ and $\beta$. Carrying out the integration in the first and second conditions given in Eq. (62) by means of the power series expansions in $E$ of the integrands, Epstein obtains

$$A_0 = \frac{2\pi^3\kappa^3\alpha}{h^2} \cdot \frac{1}{(n_1 + n_2 + n_3)^3},$$

$$\Delta A = \frac{3h^3E}{8\pi^2\kappa\alpha m}(n_1 - n_2)(n_1 + n_2 + n_3).$$

Applying Bohr's frequency relation, he then gets for the splitting of a spectral line $\Delta \nu$ in the Stark effect the formula

$$\Delta \nu = (3h/8\pi^2\kappa\alpha m)E Z,$$

$$Z = \{(m_1 - m_2)(m_1 + m_2 + m_3) - (n_1 - n_2)(n_1 + n_2 + n_3)\}.$$

The splitting of $H_\alpha$, $H_\beta$, $H_\gamma$ and $H_\delta$ lines of hydrogen are obtained by putting $\kappa = 1$, $n_1 + n_2 + n_3 = 2$, and $m_1 + m_2 + m_3 = 3, 4, 5$ and $6$. In the comparison of the corresponding differences in wavelength $\Delta \lambda$ with the experimental values,\(^{25}\) Epstein uses "Sommerfeld's selection rule", that is, $m_1 \geq n_1, m_2 \geq n_2$

and \( m_2 \geq n_3 \). As a result, he finds that it holds without exception that “an even difference \( m_3 - n_3 \) leads to parallel (\( p \)-) polarization, and an odd difference to perpendicular (\( s \)-) polarization”. \( p \)- or \( s \)-polarization corresponds respectively to \( p \)- or \( s \)-component in Stark’s paper, namely to the case in which the electric field of light is parallel or perpendicular to the external field, respectively. Epstein notes that this rule results from the standpoint “to discover lawfulness by examining available observational materials”, and that “we are at present not able to give an explanation of these circumstances and point out only that the aerial constant \( p_\varphi \) in the plane perpendicular to the direction of the field \( E \) is given by the quantum number \( n_3 \)”.

About the intensities of light, Epstein sees that “any component determined by the six quantum numbers is \( \text{caeteris paribus} \) stronger, accordingly as the largest of the three differences \( m_1 - n_2, m_2 - n_2 \) and \( m_3 - n_3 \) (‘quantum jumps’) is larger”. There are, however, cases exceptional to this. Furthermore, there are such observed lines that could only be obtained by putting \( m_3 - n_3 = -1 \). Thus, understanding Sommerfeld’s selection rule “as a matter of probability”, Epstein considers that “though its violation is quite improbable, favorable subconditions may possibly be added”, and modifies Sommerfeld’s rule in the following way: “\( m_1 \geq n_1, m_2 \geq n_2, m_3 + 1 \geq n_3 \) with the subcondition that \( m_3 + 1 = n_3 \) is allowed only when one of the differences \( m_1 - n_1 \) or \( m_2 - n_2 \) is particularly large”.

Epstein’s modification as well as Sommerfeld’s original form of the selection rule was phenomenological, as they were conscious of this feature. For the interpretation of it, besides the quantum condition some additional theory was necessary that would give the intensity of spectral lines, as will be discussed in the following sections of the present chapter.

### §6. The Stark effect and Planck’s quantum condition

In his paper “On Quantum Hypothesis”,\(^{26}\) Schwartzschild uses “the elliptic coordinates” \((\rho, \mu, \varphi)\) defined by\(^{27}\)

\[
\rho = \frac{r + z}{2}, \quad \mu = \frac{r - z}{2}; \quad r^2 = x^2 + y^2 + z^2.
\]


\(^{27}\)Though Schwartzschild calls these coordinates elliptic coordinates, they are a kind of the parabolic coordinates \((\rho = \xi^2, \mu = \eta^2 / 2)\). But, the linear, polar and parabolic coordinates can each be regarded as a special case of the elliptic coordinates.
The external electric field is in the direction of the z-axis. Putting Hamilton’s characteristic function $W$ in the form $W = W(p) + W_2(\varphi) + W_3(\varphi)$, he makes the separation of variables and obtains for the energy of the electron in the hydrogen atom $F$ in first order approximation of the strength of the external electric field $E$,

$$-F = \frac{2\pi^2 me^4}{\hbar^2} \frac{4}{(n_1' + n_2')^2} + \frac{3\hbar^2 E (n_1'^2 + n_2'^2)}{8\pi me}.$$  \hspace{1cm} (68)

The quantum numbers $n_1'$ and $n_2'$ in this equation correspond respectively to $2n_1 + n_3$ and $2n_2 + n_3$ in Epstein’s case. From Eq. (67) Schwartschild derives $\Delta \nu$ corresponding to that given in Eq. (66). He shows that the values of $\Delta \nu$ calculated by applying Sommerfeld’s selection rule $m' \geq n_1$ and $m_2' \geq n_2'$, are in fairly good agreement with those observed by Stark. He gives no discussion of the polarization and intensity of the spectral lines.

The reason why Schwartschild has arrived at the formula for $F$ which depends only on the two quantum numbers is that in his calculation use has been made of the property of $F$ being a function of $2n_1 + n_3$ and $2n_2 + n_3$. He has started from Planck’s quantum condition, differently from Epstein’s case in which Sommerfeld’s quantum condition is applied directly to $p_\xi = \partial W/\partial \xi$ etc. calculated from $W$.

Planck’s quantum condition just mentioned above is a generalization, to the case of multi-dimensional phase space, of his hypothesis of the quantum of action which we have described in Sec. 1, §3, Chap. 2 of the present volume. In his paper “The Quantum Hypothesis for Molecules with Many Degree of Freedom”, Planck considers that in order to “characterize in statistical meaning” a state of a large number $N$ of molecules each having $f$ degrees of freedom, it is necessary first to divide the whole phase space of $2f$ dimensions into elementary domains of identical probability, and thereby determine the number of molecules in each elementary domain. Planck says that “it is now characteristic of the quantum hypothesis, contrary to the classical theory, that the elementary domains of probability posses quite determined forms and sizes”.

Supposing equi-energy surfaces $g = 0, g_1, g_2, \ldots, g_n, \ldots$ in the phase space, where $g$ means the energy of a molecule, Planck considers that the elementary domains are surrounded by these surfaces. He assumes that $g_n$ is determined

---

by imposing on the $2f$-dimensional volume surrounded by the surface $g = g_n$
the condition

$$\int \int_{g=g_n} \cdots dq_1 dq_2 \cdots dq_f dp_1 dp_2 \cdots dp_f = (nh)^f. \tag{69}$$

Denoting the sizes of the elementary domains by $G_0, G_1, G_2, \ldots$ in succession, he then gets

$$G_n = \{(n+1)^f - n^f\}G_0; \quad G_0 = h^f. \tag{70}$$

$G_n/G_0$ gives the probability that a molecule comes in the $n$th elementary domain. Further, he considers that the indication of the energy $g$ does not suffice, and introduces other quantities $g', g'', \ldots$, supposing

$$\int \int_{0,0,0,\ldots} \cdots dq_1 dq_2 \cdots dq_f dp_1 dp_2 \cdots dp_f = (nh)^i (qh)^i' (rh)^i'' \cdots \tag{69'}$$

where $q, r, \ldots$ are integers and $i+i'+i''+\cdots = f$. Because the integers $n, q, r \ldots$
are independent of each other, for Eq. (69') to be satisfied, $g_n, g'_q, g''_r, \ldots$ should
be chosen adequately so that the left-hand side of Eq. (69') becomes $g_n g'_q g''_r \cdots$. From this condition he concludes that

$$g_n = (ng)^i, \quad g'_q = (qh)^i', \quad g''_r = (rh)^i'', \ldots. \tag{71}$$

Planck applies Eq. (69) to a rotator of $f = 2$, and discusses the specific heat
of hydrogen gas. He also applies Eqs. (69) and (69') to a three dimensional harmonic oscillator, taking for $g'$ the angular momentum of the oscillator ($i = 1, i' = 2$), and treats the specific heat of an aggregate of the oscillators.

Planck's hypothesis of the quantum of action was proposed in 1911 to interpret the quantum nature of oscillators emitting thermal radiation. Its generalization under consideration to systems with many degrees of freedom was done also from the standpoint of the statistical heat theory. This is in sharp contrast to the fact that Sommerfeld proposed in 1911 his hypothesis of the quantum of action to interpret molecular elementary processes, and treated in 1915 the spectral lines of the hydrogen atom by formulating his quantum condition. Planck applied his method to the problem of the hydrogen atom,\textsuperscript{30} after Sommerfeld derived the Balmer formula.

\textsuperscript{30}M. Planck, Ann. d. Phys. 50 (1916), 385.
Now, let us return to Schwartzschild's paper in which Planck's method is applied to conditional periodic systems. Citing Charlier's book,\textsuperscript{31}) which is referred to also in Epstein's paper, Schwartzschild defines a conditional periodic system to be a system, for which the solution to the Hamilton equation of motion is given in the form

\[
\begin{align*}
q_i &= q_i(\alpha_1, \alpha_2, \ldots, \alpha_f, w_1, w_2, \ldots, w_f) \\
p_i &= p_i(\alpha_1, \alpha_2, \ldots, \alpha_f, w_1, w_2, \ldots, w_f)
\end{align*}
\]

(72)

Here, the $\alpha_j$'s are constants ($j = 1, 2, \ldots, f$), and the $w_j$'s are variables called angular variables, each of which increases in proportion to the time $t$ as

\[ w_j = \nu_j t + \beta_j , \]

(73)

where $\beta_j$ is the initial value of $w_j$, and $\nu_j$ is called the "average motion". $\alpha_j$ could be taken as the canonical variable conjugate to $w_j$. In this case, the Hamiltonian, namely the energy of the system $F$, is known from $\frac{\partial F}{\partial w_j} = 0$ to be a function only of the $\alpha_j$'s, $F(\alpha_1, \alpha_2, \ldots, \alpha_f)$, while $\nu_j$ is given by the relation $\nu_j = \frac{\partial F}{\partial \alpha_j}$. $\alpha_j$ in this case is called the "action variable" conjugate to the angular variable $w_j$.

The volume of the phase space is given by

\[
\int \cdots \int d\alpha_1 dw_1 d\alpha_2 dw_2 \cdots d\alpha_f dw_f .
\]

Schwartzschild considers that "there results the following natural partition of the phase space", and puts

\[
\int d\alpha_j dw_j = 2\pi \int d\alpha_j = h .
\]

(74)

From its correspondence to Planck's elementary domain, Schwartzschild concludes that

\[
\alpha_j = \varepsilon_j + n_j \frac{h}{2\pi} ,
\]

(75)

where $\varepsilon_j$ is a constant and $n_j$ an integer.

He considers the case where Hamilton's characteristic function $W$ is obtained. From the Hamilton–Jacobi partial differential equation, in the following form in terms of the integration constants $\alpha_j$'s, the auxiliary variables $\eta_j$'s

\textsuperscript{31}) C. L. Charlier, Ref. 24) on p. 134 in the present chapter.
and a function \( V(\alpha_j, \eta_j) \):

\[
W = \alpha_1 \eta_1 + \alpha_2 \eta_2 + \cdots + \alpha_f \eta_f + V(\alpha_j, \eta_j),
\]

where \( V(\alpha_j, \eta_j) \) as well as the \( q_i \)'s is supposed to be represented by periodic functions of \( \eta_j \)'s of period \( 2\pi \). In this case, from the relations \( p_i = \partial W/\partial q_i \) and \( w_j = \partial W/\partial \alpha_j \), it is known that the \( \eta_j \)'s, and hence the \( p_i \)'s as well as the \( q_i \)'s, are functions of the \( w_j \)'s of period \( 2\pi \). This shows that the \( \alpha_j \)'s are action variables and the \( w_j \)'s are angular variables. Schwartzschild applies this method to the calculation of the Stark effect mentioned above, with the use of the curvilinear coordinates \( q_1 = \rho, q_2 = \mu \) and \( q_3 = \varphi \) given by Eq. (67). He treats also by this method a rotator to discuss molecular band spectra.

Schwartzschild notes that his “partition coincides in many cases, but not in all, with the partitions proposed by Planck and Sommerfeld”. It is shown by Epstein in a paper in 1918 that Planck's quantum condition applied by Schwartzschild to conditional periodic motions is equivalent to Sommerfeld’s quantum condition.

Epstein starts his paper “On the Structure of the Phase Space of Conditional Periodic System”\(^{32}\) from the fact that for a conditional periodic motion \( p_i \) can be put in the form

\[
p_i = p_i(q_i; \alpha_1, \alpha_2, \ldots, \alpha_f),
\]

(76)

if the coordinates are properly chosen. In this expression, the \( \alpha_j \)'s are constants and \( p_i \) does not depend on the \( q_j \)'s other than \( q_i \). In the case of a libration, \( q_i \) has terminal points so that one has \( a_i \leq q_i \leq b_i \). In the case of a rotational motion, \( q_i \) becomes a cyclic coordinate denoted by \( q_z \). \( q_z \) has no terminal point, but the motion is periodic and \( p_z \) is constant.

On the basis of Sommerfeld’s quantum condition, Epstein puts

\[
2 \int_{a_i}^{b_i} p_i dq_i = n_i h \quad \text{or} \quad \int_{0}^{2\pi} p_z dq_z = 2\pi p_z = n_z h,
\]

(77)

Because the \( f \) variables \( u_i \)'s defined by

\[
u_i = 2 \int_{a_i}^{b_i} p_i dq_i / 2\pi \quad \text{or} \quad u_z = \int_{0}^{2\pi} p_z dq_z / 2\pi,
\]

(78)

can be determined by the \( f \) variables \( \alpha_i \)'s, the \( u_i \)'s may be used instead of the \( \alpha_i \)'s to give

\[
p_i = p_i(q_i; u_1, u_2, \ldots, u_f). \quad (79)
\]

Solving this set of equations for the \( u_i \)'s one gets

\[
u_i = u_i(p_1, p_2, \ldots, p_f; q_1, q_2, \ldots, q_f). \quad (80)
\]

Equation (80) gives \( f \) surfaces in the phase space. The volume of the region in the phase space, which is surrounded by the surface specified by the \( u_i \)'s and \( (u_i + du_i) \)'s, is given by

\[
dG = dG = \int_{u_i}^{u_i+du_i} \cdots \int dp_1dp_2\cdots dp_fdq_1dq_2\cdots dq_f. \quad (81)
\]

Transforming the integration variables into the \( u_i \)'s and \( q_i \)'s with the use of Eq. (79), one gets

\[
dG = du_1du_2\cdots du_f \int \cdots \int \left| \frac{\partial p_i}{\partial u_j} \right| dq_1dq_2\cdots dq_f.
\]

The \( i-j \) element \( \partial p_i/\partial u_i \) of the determinant \( \left| \partial p_i/\partial u_j \right| \) depends only on \( q_i \) on account of Eq. (79). Transforming the integral with the use of the derivative of Eq. (78), and noting that \( |\partial u_i/\partial u_j| = 1 \), one finally gets

\[
\begin{cases}
dG = dq_1dq_1\cdots dq_f, \\
dq_i = 2\pi du_i \quad \text{or} \quad dq_z = 2\pi du_z,
\end{cases} \quad (82)
\]

which, together with Eqs. (77) and (78) show the equivalence of Sommerfeld's quantum condition with Planck's.

§7. The quantum condition and the Zeeman effect

The quantum condition was applied to the calculation of the Zeeman effect in 1916 by Debye and Sommerfeld. In the paper "Quantum Hypothesis and Zeeman Effect",\(^{33}\) Debye cites the papers on the Stark effect of Epstein and Schwartzschild mentioned above, and puts \( W_i = \int p_i dq_i = n_i\hbar \). \( W_i \) is obtained by integrating \( p_i = dW_i/dq_i \) which results from the separation of variables of the Hamilton–Jacobi equation for Hamilton's principal function \( S \), with the

supposition \( S = \sum_i W_i(q_i) - \alpha t \). The external magnetic field of strength \( H_0 \)
is assumed to be in the direction the \( z \)-axis. The Lagrangian is given by

\[
L = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{eE}{r} + \frac{eH_0}{2c} (xy^2 - yx^2),
\]

where \( E \) denotes the electric charge of the nucleus. The separation of variables
of the Hamilton–Jacobi differential equation obtained from this Lagrangian is
made with the use of the spherical coordinates \((r, \theta, \varphi')\) rotating round the
\( z \)-axis. \( \varphi' \) is related to the azimuthal angle \( \varphi \) by

\[
\varphi' = \varphi - \omega t; \quad \omega = eH_0/2mc.
\]

Debye applies the quantum condition given above by putting \( q_1 = r, q_2 = \theta, \)
\( q_3 = \varphi' \), and derives the energy \( \alpha \) of the system

\[
\alpha = -\frac{2\pi^2 me^2 E^2}{h^2 \left( n_1 + n_2 - n_3 \right)^2} + n_3 \hbar \frac{\omega}{2\pi}.
\]

Because, on this quantum condition, the square of the angular momentum of
the electron becomes \((n_2 - n_3)^2(h/2\pi)\), he defines \((n_1, n_2 - n_3, n_3)\) newly as
\((n_1, n_2, n_3)\), so that \( \alpha \) is rewritten as

\[
\alpha = -\frac{2\pi^2 me^2 E^2}{h^2} \frac{1}{(n_1 + n_2)^2} + n_3 \hbar \frac{\omega}{2\pi}.
\]

Making use of Bohr’s frequency relation, he then gets for the splitting of a
spectral line in the Zeeman effect the formula

\[
\Delta \nu = (m_2 - n_3) \frac{\omega}{2\pi}.
\]

Debye says that “the ordinary triplet arises when one allows only the smallest
integer numbers 0 and 1 for \( m_3 \) and \( n_3 \)”, but does not go further into the
multiplicity of the splitting.

Sommerfeld treats the Zeeman effect in his paper “On the Theory of the
Zeeman Effect of the Hydrogen Lines”, \(^{34}\) by solving directly the Newton equation
of motion for the electron in a hydrogen atom. (He uses, however, the
Hamilton–Jacobi partial differential equation in his discussion of relativistic
effects on the Zeeman effect.) Making use of the rotating spherical coordinates

\(^{34}\) A. Sommerfeld, Phys. Zeit. 17 (1916), 491.
\( (r, \theta, \varphi') \) as in Debye’s work, one gets the same type of equation of motion as that in the case of no external motion if one neglects terms of the order of \( \omega^2 \), as we have seen in Larmor’s theory in Sec. 3, §6, Chap. 2, Vol. I. Denoting the radial, azimuthal and directional quantum numbers in this system of coordinates with \( n', n_1 \) and \( n_2 \), respectively, one thus has for the energy of the electron in this system of coordinates

\[
W' = -\frac{\mathcal{R}h}{(n+n')^2}; \quad n = n_1 + n_2.
\]

Because the energy in the spherical coordinates at rest is given by

\[
W = W' + \omega p_{\varphi'},
\]

neglecting again \( \omega^2 \) one obtains, with the use of

\[
p_{\varphi'} = n_1 h/2\pi,
\]

\[
\Delta \nu = \left(\frac{\omega}{2\pi}\right) Z; \quad Z = m_1 - n_1.
\]  \hspace{1cm} (87)

Sommerfeld says that “our result does not become coincident with the conclusion of the classical theory, but becomes close to it”. For \( H_\alpha \), Sommerfeld’s selection rule \( m_1 \geq n_1 \) gives \( Z = 0 \) and \( 1 \) because \( m + m' = 3 \) and \( n + n' = 2 \). From this he concludes that \( H_\alpha \) is triplet, by arguing that the line corresponding to \( Z = -1 \) is equivalent to that corresponding to \( Z = 1 \), for the reason that the electron can rotate in two senses round the direction of the magnetic field as axis. For \( H_\beta \), there can also be the case \( Z = 2 \), and for \( H_\gamma \), the case \( Z = 3 \). In general, for \( M = m + m' > N = n + n' \), the maximum value of \( Z \) is \( M - N + 1 \), which does not necessarily give rise to triplet. But, he points out that the rule shown by Epstein for the Stark effect, that an even (odd) jump in the azimuthal quantum number brings on polarization parallel (perpendicular) to the external field, is effective also for the Zeeman effect, especially for the cases of \( Z = 0 \) and \( Z = 1 \).

About the result coming from the quantum condition that shows the existence of extra components other than triplet ones, Sommerfeld notes that “an older theory of Ritz has given rise to surplus components of the same situation”. He says that “it is right improbable that these components might be real and unobserved only because of their weaker intensity”, suggesting that, here also, the combination rule could not be unconditionally applied. On the other hand, Sommerfeld mentions that “even if the conclusion is by no means satisfactory, the power of the quantum theory in conjunction with the Bohr model nevertheless surprises us also in this case”. Sommerfeld mentions further as follows — “This power goes even beyond the expectation that Mr. Bohr
himself has expressed with regard to the Zeeman effect. He thought that the electric splitting of the spectral lines would be brought about by a change of the orbit of the electron and its energy, while the magnetic splitting by a magnetic influence on the transition from the initial to final orbits, which would result in a change of the frequency rule” —.

Bohr mentioned this idea in his paper “On the effect of Electric and Magnetic Fields on Spectral lines”, though in Sommerfeld’s paper no explicit reference to Bohr’s paper is given. In this paper, Bohr calculated the energy $A$ of electron changed by taking into a weak electric field with the use of the ordinary dynamics, by consideration the slight distortion of its orbit by the electric field. On the assumption that for the “transition” from the $(n + 1)$th state to the $n$th state, $(A_{n+1} - A_n)/h$ should approach $(dA_n/dn)/h$ in the case of large $n$, as was already supposed in his paper on the Bohr model, he calculated $A_n$ by integrating $dA_n/dn$ with respect to $n$. In the case of the Zeeman effect, however, according to the ordinary dynamics the electron continues making a circular motion, so that he could not apply the method of calculation used for the Stark effect. Bohr thus supposed that the frequency-energy relation “is replaced by” a relation including the strength of the magnetic field. On the contrary, Sommerfeld concludes that “with literal preservation of this relation there results a change in the energy of the orbit of the electron, which leads to the principal point of the Lorentz triplet, without any new assumption”, appreciating the progress of the quantum theory because of the quantum condition.

Importance should be attached to the fact that, while the application of the quantum condition to the Zeeman effect introduced the Planck constant $h$, it produced the Larmor frequency which bears no relation to $h$ for the splitting in frequency of spectral lines, as in the classical theory. As has been seen in Sec. 1, §3, Chap. 2, Vol. I, the specific electric charge $e/m$ was known for the first time, from experimental results on the Zeeman effect by means of the Larmor frequency. This established the first existential or substantial, irrevocably to $h$, foundation for the theory of electron. The same situation held also for the experimental confirmation of the Nagaoka model of the atom with nucleus by means of the Rutherford scattering, which is independent of $h$ even in quantum mechanics. In other words, those classical theoretical laws that were later found to be valid also in the quantum theory gave openings to the world of the quantum theory.

35) N. Bohr, Phil. Mag. 27 (1914), 506.
3. The Transition Probability

§1. Two trends of thinking — energetics and atomism

We have seen, in §6 of the preceding section, that Planck generalized in 1915 his hypothesis of the quantum of action to systems of many degrees of freedom. This generalization arose from a question made by Poincaré to Planck at the 1st Solvay Conference held in 1911. There Planck proposed, as the principle to give the energy \( n h \nu \) of a one-dimensional oscillator, the partition of its phase space into the elemental domains of magnitude \( \int dpdq = h \). Poincaré asked the following\(^{36}\) — "As regards the elementary quantity of action, it must undoubtedly be represented by \( \int dpdq \), if only a single degree of freedom is concerned, and by

\[
\int \int (dp_1dq_1 + dp_2dq_2 + \cdots + dp_ndq_n),
\]

when about it \( n \) are available. One is led in this way, for example in the case of a single degree of freedom, to cut up the \( pq \)-plane into elements of area \( h \); since this cutting up is possible in various ways, one can divide it for example into squares or into any other arbitrary figures. Does the way of this cutting up have any effect on the results?"

In a footnote attached to the proceeding, Poincaré added — "The phase space of Gibbs consists of the \( 2n \)-fold integral

\[
\int dp_1dq_1dp_2dq_2\cdots dp_ndq_n,
\]

but this would no longer possess the same dimensions as \( h \), that is, energy multiplied by time. The double integral considered here possesses the same character as the phase space, that is, it remains invariant by the Hamilton equations." Shortly after in the course of discussion at the Conference, Poincaré again asked his question, saying "Now, may there be many degrees of freedom?" Considering a three-dimensional oscillator, one should have \( a h \nu \) for the energy of its motion along the \( x \)-axis, with an integer \( a \). For the \( y \)- and \( z \)-axes one should have similarly the energies \( \beta h \nu \) and \( \gamma h \nu \), with integers \( \beta \) and \( \gamma \). Taking another system of axes, one would have other energies, \( \alpha' h \nu, \beta' h \nu, \gamma' h \nu \), say,

The Transition Probability

“α’, β’, γ’ should be integers, even if one would arbitrarily select the new axes; and this is not possible.” 37

Planck mentioned in the discussion “The division of the plane in ellipses is not arbitrary, but is necessary for the calculation of the probability of a given amount of energy of the oscillator. If a calculation of the probability that another quantity shall have a determined value is concerned, one must naturally introduce another division.” Planck answered is concerned that “for many degrees of freedom there is not yet any quantum theory”, and “nevertheless, I think it not at all impossible to establish such a theory”. For Planck it was above all important to treat energy, even in the meaning of the statistical heat theory.

In his lecture itself, 38 which was made in the 1st Solvay Conference to propose his hypothesis of the quantum of action, he said that “because the intensity of the black body radiation is, according to Kirchhoff, independent of the emitting and absorbing substance, one will be allowed to expect that a radiator provided with ideal characters presents the correct value of the black body radiation, unless its behaviour is in contradiction with only the principles of the dynamics and electrodynamics”. He treated thus a charged linear oscillator as the simplest electromagnetic radiator. He expressed the energy of the resonator as given in Eq. (11) in the preceding chapter. Planck’s resonator was not explicitly given any mass or charge, and he stressed its connection with energy, as we have mentioned in Sec. 1, §1, Chap. 1 of the present volume. As we have pointed out there also, Planck’s resonator was introduced into his theory as a thing that was (on the average over an aggregate) provided with suitable thermodynamical attributes abstracted from the functions of the atom in Michelson’s or Wien’s theory.

As Planck took the standpoint of the thermodynamics, namely that of energetics, he went the way of applying the 2n-fold integral in Gibbs’ phase space, referred to by Poincaré, in generalizing his hypothesis of the quantum of action to molecular systems with many degrees of freedom. Planck connected the 2n-fold integral, however, not directly to $\hbar$, but to $h^n$ as Poincaré pointed out. As we have seen in the preceding section, Planck’s generalization of the hypothesis of the quantum of action was done at first in relation to the theory of specific heat.

38) Cf. Sec. 1, §3, p. 46, Chap. 2 of the present volume.
The double integral described by Poincaré is connected, as Poincaré said, with the Hamilton equation of motion in dynamics. In fact, if one expresses $\int \int dp_idq_i$ as $\int p_idq_i$, by considering $p_i$ to be a function of $q_i$ according to Sommerfeld, the double integral described by Poincaré becomes $\sum \int p_idq_i$, which corresponds to $\sum W_i(q_i)$ in the formulation due to Epstein and Schwartzschild, that is, to Hamilton’s characteristic function in the form with separation of the variables. Ishiwara’s quantum condition may be understood to be in this way. Poincaré’s question to Planck shows that Poincaré took Planck’s hypothesis of the quantum of action to be of dynamical nature. Poincaré mentioned, however, in his footnote we have quoted above, of its thermodynamical nature. The term “statistical orbit” used by Epstein seems to be somewhat Planck-like.

Sommerfeld’s quantum condition was proposed from the standpoint of atomism, contrarily to Planck’s standpoint. When Sommerfeld gave his lecture on “the meaning of the quantum of action for aperiodic molecular processes in physics”, at the 1st Solvay Conference at the same time as Planck, he treated the problem of “a single physical event”, or “pure molecular processes”, namely, elementary processes in the later terminology, and analyzed the production of X-rays by cathode rays, the photoelectric effect and the potential difference for ionization, by means of the action integral. Nicholson, who derived suggestions from Sommerfeld’s’ viewpoint shown at the 1st Solvay Conference, arrived at the concept of the quantum nature of angular momentum. This concept was accepted in the Bohr model, revealing that the angular momentum of an electron rotating in a circular orbit is an integer multiple of $h/2\pi$.

Sommerfeld grasped the quantum nature of angular momentum in the unit of $h/2\pi$ in its connection with the action integral, and imposed the quantum condition $\int pdq = nh$ on the rotational motion of an electron in a Kepler orbit. Not having succeeded in arriving at the Balmer formula with only this method, however, he was driven by the necessity of quantizing the eccentricity of the Kepler orbit, and imposed the quantum condition also to the radial motion. By this procedure, Sommerfeld’s quantum condition was generalized to motions with many degrees of freedom, and the directional quantization was derived.

The two fundamental assumptions in the Bohr model, namely, the existence of discrete energy states in an atom and the energy-frequency relation for spectral lines, were strengthened by Sommerfeld’s quantum condition. Now, being given these facts, Einstein introduced in 1916 the transition probability for a transition between discrete states of an atom, by the mediation of which
he attempted to understand the Planck formula of black body radiation in a statistical-theoretic way, as we shall see in the next subsection.

\section{Understanding of the radiation formula in terms of transition probabilities}

In his 1916 paper “Emission and Absorption of Radiation according to the Quantum Theory”,\textsuperscript{39} Einstein starts his discussion from the problem of the relation between the average energy of a resonator $E_\nu$ and the energy density of radiation $\rho_\nu$, namely

\begin{equation}
E_\nu = \frac{c^3}{8\pi\nu^2} \rho_\nu.
\end{equation}

As we have seen in Chap. 1 of the present volume, this relation gives the Planck formula for $\rho_\nu$ when the value of $E_\nu$ for Planck’s resonator is used, although this relation is derived by means of electromagnetics. Einstein says that “it remains however unsatisfactory that the electromagnetic-mechanical consideration, which leads to Eq. (88), is not compatible with the basic idea of the quantum theory”.\textsuperscript{40}

On the other hand, he considers that “since the Bohr theory of spectra has achieved great success, it seems not to be doubtful that the basic idea of the quantum theory must be held fast”. He thus raises the problem that “the electromagnetic-mechanical considerations, which led Planck to Eq. (88), must be replaced by quantum-theoretical considerations of the interaction between matter and radiation”.\textsuperscript{40}

Now, the Brownian motion was one of Einstein’s themes of study since 1905, and his way of thinking about this theme was brought into his works that we have treated in Vol. I. In his paper under consideration, also, he says that “the behaviour of a monochromatic resonator in radiation field in accordance with the classical theory can be surveyed by the following way of consideration used for the first time in the theory of the Brownian motion”. Let $E$ be the energy of the resonator at a time. The change in $E$ after a time duration $\tau$, which is supposed to be long compared with the period of the resonator but so short that the relative change in $E$ can be treated as infinitesimal, consists of two kinds of the change in $E$, $\Delta E_1$ and $\Delta E_2$, say.

\textsuperscript{40}In the original paper, Eq. (88) is numbered as (1).
ΔE₁ is the change caused by "out-radiation" (Ausstrahlung), namely by the emission of radiation, and is given by

$$ΔE₁ = -AEτ.$$

ΔE₂ comes from the work exerted by the electric field on the resonator. This is proportional to the radiation density ρ, and has value and sign depending on "chance". Einstein names it change due to "in-radiation" (Einstrahlung). From an electromagnetic-statistical consideration its average value is given by

$$\overline{ΔE₂} = Bρτ.$$

Because the average value of E taken over a large number of the resonators should be independent of time, it must hold that

$$E + ΔE₁ + ΔE₂ = E̅,$$

or

$$E̅ = (B/A)ρ.$$

If one calculates the constants A and B for the monochromatic resonator in the usual way by means of electromagnetics and mechanics, one is led to Eq. (88). Einstein says that "we will now carry out a corresponding calculation on a quantum-theoretical basis without any assumption to specialize the things in interaction with the radiation, which will be named 'molecules'".

Consider a gas of identical molecules, which is in statistical equilibrium with thermal radiation. Each molecule is supposed to be only in a discrete series of state Z₁, Z₂, ..., having the energies ε₁, ε₂, ..., respectively. From the Boltzmann principle, the probability of the state Zₙ, that is, the relative number Wₙ of the molecules which stay in the state Zₙ, is

$$Wₙ = pₙe^{-εₙ/kT}, \quad (89)$$

where k is the Boltzmann constant, and pₙ is the "weight" of Zₙ, namely, a constant characteristic of the respective quantum state of the molecule, which is independent of the temperature T.

Einstein assumes that a molecule can transit from the state Zₙ to the state Zₘ by absorbing radiation of a definite frequency υₙₘ, as well as from the state Zₘ to Zₙ by emitting radiation of the same kind. The energy of radiation thereby changed by εₘ - εₙ (εₘ > εₙ). "In thermal equilibrium
there must exist statistical equilibrium about every one of these elementary processes. We can therefore restrict ourselves to the consideration of a single elementary process belonging to a pair of determined indexes \((n, m)\)."

For these elementary processes Einstein sets up "simple hypotheses, by which we are led through the limit to the classical theory shortly sketched above". In his following paper "On the Quantum Theory of Radiation", \(^{41}\) Einstein writes that "of the laws which are determinative of these transitions, we introduce certain hypotheses, which are obtained by the translation of the relations known for a Planck’s resonator in accordance with the classical theory to the yet unknown ones of the quantum theory”.

Now, in "out-radiation", corresponding to emission of energy from a resonator without being excited by external field, "the transition will take place without external influence". Einstein says that "one can hardly think of it other than according to the way of radioactive reaction", and puts forth as the first assumption that the number of transitions per unit time is

\[
A_n^m N_m,
\]

where \(N_m\) is the number of the molecules in the state \(Z_m\), and \(A_n^m\) is a constant proper to the combination of \(Z_m\) and \(Z_n\). In the paper "On the Quantum Theory Radiation", \(^{41}\) he writes — "The assumed statistical law corresponds to that of a radioactive reaction, and the assumed elementary process to the reaction in which only \(\gamma\)-rays are emitted. It needs not be assumed that this process demands no time; this time must only be negligible compared with the time during which the molecule is in the state \(Z_1\), etc."

The "in-radiation" is brought about by the radiation in which the molecule stays. In the case of a radiator, "it can cause an increase as well as a decrease in energy". In our case, Einstein therefore considers that "it can cause similarly a transition \(Z_n \rightarrow Z_m\) as well as a transition \(Z_m \rightarrow Z_n\)", \(^{42}\) and supposes as the


\(^{42}\) The "in-radiation" includes, in this way, absorption as well as induced emission of radiation in present-day terminology. On the processes \(Z_n \rightarrow Z_m\) and \(Z_m \rightarrow Z_n\), Einstein wrote in the paper cited in Ref. 41) "We name both the events 'change of state by in-radiation'", and called them in 1922 the "positive and negative in-radiations" [cf. Footnote 34) in Sec. 1, §4, Chap. 4, Vol. II]. The "out-radiation" is the spontaneous emission in present-day terminology. It may be noted that the German word "Ausstrahlung" corresponds originally to emission of radiation in English, and "Einstrahlung" to irradiation (= in + radiation). We prefer to use "out-radiation" and "in-radiation", because of the contrast between "aus" and "ein" and of their characteristic correspondence with the present-day terms mentioned above.
second assumption that the number per unit time of the transition $Z_n \rightarrow Z_m$ is expressed by

$$B^m_n N_n \rho,$$

and that of the transition $Z_m \rightarrow Z_n$ by

$$B^m_n N_n \rho,$$

where $B^m_n$ and $B^m_m$ are constants related to the combination of $Z_n$ and $Z_m$.

From the condition for statistical equilibrium, it follows that

$$A^m_m N_m + B^m_m N_m \rho = B^m_n N_n \rho.$$  \hfill (90)

On the other hand, Eq. (89) shows that

$$\frac{N_n}{N_m} = \left( \frac{p_n}{p_m} \right) e^{(\epsilon_m - \epsilon_n)/kT}. \hfill (91)$$

From these equations follows the relation

$$A^m_m p_m = \rho \{ B^m_n p_n e^{(\epsilon_m - \epsilon_n)/kT} - B^m_m p_m \}. \hfill (92)$$

If $\rho$ is supposed to increase indefinitely with $T$ as the third assumption, it could be deduced that

$$B^m_n p_n = B^m_m p_m,$$  \hfill (93)

It results therefore that

$$\rho = \frac{\alpha_{mn}}{e^{(\epsilon_m - \epsilon_n)/kT} - 1}; \quad \alpha_{mn} = \frac{A^m_m}{B^m_m}. \hfill (94)$$

This is the Planck relation with constants not being determined. The constants $A^m_m$ and $B^m_m$ would be directly calculated, if we are in possession of an electrodynamics and mechanics modified in the meaning of the quantum hypothesis.” — Einstein concludes so. In this way, the Planck radiation formula is derived, from such three assumptions as have “simplicity” and “generality with which considerations are carried through unrestrictedly”. Einstein mentions that “it seems to me to be very probable that they will be the main features of future theoretical description”. It will be seen in the following sections that this expectation proves true.
He points out that from Wien's displacement law one has $\alpha_{mn} \propto \nu^3$ and $\varepsilon_m - \varepsilon_n \propto \nu$, so that one gets, with a universal constant $h$,

$$\varepsilon_m - \varepsilon_n = h\nu.$$  \hspace{1cm} (95)

In the paper "On the Quantum Theory of Radiation", Einstein writes that Eq. (95) "constitutes, as is well known, the second principle in Bohr's theory of spectra, of which one is already allowed from Sommerfeld's and Epstein's completion to assert that it belongs to a secured existence in our science".

In this paper, also, he says that "the formal similarity of color distribution curve of thermal radiation to Maxwell's law of velocity distribution is very remarkable", and "through this similarity W. Wien was already led to a far-reaching determination of the radiation formula", but, as the quantum theory developed, "that consideration of Wien, which has led to Wien's radiation formula, sunk already of course into oblivion". These circumstances indicated by Einstein are just those which we have described in Sec. 1, Chap. 1 of the present volume. Einstein also mentions that "shortly before, I found such a derivation of the Planck radiation formula as is similar to Wien's original consideration and supported by the fundamental assumption of the quantum theory".

The method of statistical mechanics, with which thermodynamical phenomena are clarified from molecular motions, was thus developed very much on account of the introduction of the concept of transition probability. A view was obtained of the way that might lead to the unification of the two trends, energetics and atomism, by the intermediation of the transition probability.

§3. Light-quantum theoretical radiation and molecular recoils

In absorption and emission of radiation, there takes place also a momentum transfer to molecules. Einstein considers that "this leads at the same time to the fact that a determined molecular velocity distribution arises only through the interaction of radiation and molecules". This distribution should clearly coincides with the Maxwell distribution. Moreover, "this must hold independently of the nature of the molecules under consideration and of the frequencies of radiation absorbed and emitted by them". It is the main purpose of Einstein's paper "On the Quantum theory of Radiation", cited in the preceding subsection, to show that "this extensive requirement will in fact be satisfied very generally".
To get the mentioned result, however, "there arises the question of whether a molecule receives an impulse or not, when it absorbs or emits the energy \( \varepsilon \)." In the classical theory, when a body radiates an energy \( \varepsilon \), it receives the recoiling impulse \( \varepsilon / c \), "if the total amount of radiation \( \varepsilon \) is radiated along the same direction". However, "if the emission arises through a spatially symmetrical event, for example a spherical wave, there results no recoiling impulse at all". Einstein says that "these alternatives play a role also in the quantum theory of radiation". These words of Einstein should be taken to mean a conflict between the wave theory and light quantum theory. He next stresses that "it is shown that we arrive at an noncontradictory theory, only when we grasp these elementary processes as completely directional events". In other words, Einstein makes the following analysis according to the light quantum-theoretical way of thinking.

In an elementary process of out-radiation \( (Z_m \rightarrow Z_n) \), a momentum of amount \( (\varepsilon_m - \varepsilon_n)/c \) is transferred to the molecule. "If the molecule is isotropic, we must assume all the directions of the out-radiation to be equally probable. If the molecule is not isotropic, we arrive at the same statement on the assumption that its orientation changes in lapse of time according to the law of chance." For an elementary process of in-radiation, the constants \( B^n_m \) and \( B^m_n \) should be independent of the direction, "by the assumption of the isotropy or pseudoisotropy (through taking time average)".

Now, Einstein takes the case in which a molecule moves in radiation with a speed \( u \) along one direction, and considers the change of the momentum of the molecule during a short time \( t \). It consists of \( -Ru \tau \) due to the frictional force which is exerted on the molecule by the radiation and \( \Delta \) due to the irregularity of the action on the molecule by the radiation. Because the square of the changed momentum should be equal to that of the original momentum, on an average over a long enough time or over a large number of molecules, one has

\[
\overline{\Delta^2} = 2RMu^2\tau,
\]

where \( M \) denotes the mass of the molecule. This equation corresponds to Eq. (35) in Sec. 3, §2, Chap. 1 of the present volume, where we have presented Einstein's work in 1909 on the fluctuation of radiation pressure. The mirror in that case corresponds to the molecule in the present case, and the fluctuation of the velocity of the mirror to that of the momentum of the molecule. In the thermal radiation of temperature \( T, Mu^2/2 \) must be equal to \( kT/2 \). One
The Transition Probability

gets therefore

\[
\frac{\Delta^2}{\tau} = 2RkT. \tag{97}
\]

As is seen below, Einstein expresses \( \Delta^2 \) and \( R \) in terms of the transition probabilities, and shows that Eq. (97) is identically satisfied if the radiation density \( \rho \) obeys the Planck formula.

Suppose that the molecule moves along the \( x \)-axis of the coordinate system \( K \), and is at rest in the coordinate system \( K' \). The radiation is assumed to be isotropic in \( K \). The amount of the radiation per unit volume, which is contained in the frequency interval \( d\nu \) and in the solid angle \( d\kappa \), is given by

\[
\rho d\nu d\kappa / 4\pi. \tag{98}
\]

In \( K' \) this becomes

\[
\rho'(\nu', \varphi') d\nu' d\kappa' / 4\pi, \tag{98'}
\]

where \( \varphi' \) and \( \psi' \) denote the zenithal and azimuthal angles, respectively, with the \( x' \)-axis as the zenithal axis. Between Eqs. (98) and (98') the same law of transformation should hold, as that between the amplitudes squared \( A^2 \) and \( A'^2 \) of the plan waves going in the respectively directions. Neglecting terms of power higher than \( u/c \), one has therefore

\[
\frac{\rho'(\nu', \varphi') d\nu' d\kappa'}{\rho(\nu) d\nu d\kappa} = 1 - 2\frac{u}{c} \cos \varphi. \tag{99}
\]

One has, furthermore,

\[
\nu' = \nu \left( 1 - \frac{u}{c} \cos^2 \varphi \right), \tag{100}
\]

\[
\cos \varphi' = \cos \varphi + \frac{u}{c} \cos^2 \varphi, \tag{101}
\]

\[
\psi' = \psi, \tag{102}
\]

to the same order of approximation, from which one gets approximately \( \nu = \nu' \{ 1 + (u/c) \cos \varphi' \} \) and therefore

\[
\rho(\nu) = \rho(\nu') + \frac{\partial \rho(\nu')}{\partial \nu} \frac{u}{c} \nu' \cos \varphi'. \tag{103}
\]
One has also the relations
\[
\frac{dv}{d\nu'} = 1 + \frac{u}{c} \cos \varphi', \\
\frac{d\kappa}{d\kappa'} = \frac{d(\cos \varphi)}{d(\cos \varphi')} = 1 - \frac{2u}{c} \cos \varphi'.
\]
Thus one obtains
\[
\rho'(\nu', \varphi') = \left[ (\rho)_{\nu'} + \frac{u}{c} \nu' \cos \varphi' \left( \frac{\partial \rho}{\partial \nu} \right)_{\nu'} \right] \left( 1 - \frac{3u}{c} \cos \varphi' \right). \tag{104}
\]

Now, because the time during which the molecule stays in the state $Z_n$ per second is
\[
p_n e^{-\varepsilon_n/kT}/S; \quad S = p_n e^{-\varepsilon_n/kT} + p_m e^{-\varepsilon_m/kT},
\]
the number of occurrences per second of the elementary process is given by
\[
p_n e^{-\varepsilon_n/kT} B_n^m \rho'(\nu', \varphi') d\kappa'/4\pi S,
\]
for the in-radiation of the type $Z_n \rightarrow Z_m$. In each of them, the momentum $(1/c)(\varepsilon_m - \varepsilon_n) \cos \varphi'$ is transferred to the molecule in the direction of the positive $x$-axis. For the in-radiation of the type $Z_m \rightarrow Z_n$, the corresponding number is given by interchanging the indexes $m$ and $n$. But in the latter case, the momentum transferred to the molecule is $-(1/c)(\varepsilon_m - \varepsilon_n) \cos \varphi'$.

By means of Eqs. (93) and (95), the momentum transferred per second to the molecule by the in-radiation becomes
\[
\frac{h\nu'}{cS} p_n B_n^m (e^{-\varepsilon_n/kT} - e^{-\varepsilon_m/kT}) \int \rho'(\nu', \varphi') \cos \varphi' \frac{d\kappa'}{4\pi}.
\]
On inserting Eq. (104) in this and carrying out the integration, the above becomes
\[
-\frac{h\nu}{c^2 S} \left( \rho - \frac{1}{3} \nu \frac{\partial \rho}{\partial \nu} \right) p_n B_n^m (e^{-\varepsilon_m/kT} - e^{-\varepsilon_n/kT}) \nu,
\]
where $\nu'$ is now denoted by $\nu$. For the elementary process of out-radiation, there is no preferential direction in the coordinate system $K'$, so that on the average no momentum is transferred to the molecule. The coefficient $R$ of the
frictional force exerted on the molecule is, therefore, given by

\[
R = \frac{\hbar \nu}{c^2 S} \left( \rho - \frac{1}{3} \nu \frac{\partial \rho}{\partial \nu} \right) p_n B_n^m e^{-\varepsilon_n/kT} (1 - e^{-\hbar \nu/kT}). \tag{105}
\]

Next, Einstein calculates \( \Delta^2 \) in the following way. If momenta \( \lambda_1, \lambda_2 \ldots \) in the direction of the x-axis are given to the molecule by a number of mutually independent events, \( \Delta \sum \lambda_\nu \). If the average value of each \( \lambda_\nu \) is zero, then \( \Delta^2 = \sum \lambda_\nu^2 \). Moreover, if every \( \lambda_\nu^2 \) is assumed to be equal to \( \bar{\lambda}^2 \), and the number of the events giving rise to the momentum transfer to the molecule is \( l \), there holds the relation

\[
\Delta^2 = l \bar{\lambda}^2. \tag{106}
\]

In each process of in-radiation, the momentum transferred to the molecule is

\[
\lambda = \left( \frac{\hbar \nu}{c} \right) \cos \varphi,
\]

where \( \varphi \) is "the angle between the x-axis and a direction selected by the law of chance". It follows therefore that

\[
\bar{\lambda}^2 = \left( \frac{1}{3} \right) \left( \frac{\hbar \nu}{c} \right)^2. \tag{107}
\]

\( l \) is the total number of the elementary processes that occur in the time \( \tau \), and is equal to two times the number of the processes \( Z_n \rightarrow Z_m \) of in-radiation. That is,

\[
l = (2/S)p_n B_n^m e^{-\varepsilon_n/kT} \rho \tau. \tag{108}
\]

From Eqs. (106)–(108), one has

\[
\frac{\Delta^2}{\tau} = \left( \frac{2}{3S} \right) \left( \frac{\hbar \nu}{c} \right)^2 p_n B_n^m e^{-\varepsilon_n/kT} \rho. \tag{109}
\]

The application of the Planck formula to Eqs. (105) and (109) shows that there holds the relation of Eq. (97) which is intended to prove.

\section*{§4. Demand for a quantum-natured theory of radiation}

Einstein, in the conclusion of his paper under consideration, states that "the result regarding the momentum transferred to the molecule by the in-radiation
and out-radiation appears most important” to him. “If the molecule is under the action of several directional bundles of radiation, only one of them is always concerned in an elementary event of in-radiation. This bundle then determines by itself the direction of the momentum to be transferred to the molecule.” “The process of out-radiation is also a directional one. The molecule suffers in the elementary process of out-radiation a recoil impulse of quantity $h\nu/c$ in a direction that is determined in the present state of the theory only by chance.”

He continues to state as follows. “The characters of the elementary processes demanded by Eq. (97) make the construction of a theory of radiation having proper quantum nature appear almost inevitable. The weakness of the theory lies, on one side, in that it does not bring us nearer to the connection with the wave theory, and on the other side, in that it entrusts time and direction of the elementary process to ‘chance’. Nevertheless I think much of a full trust in the reliability of the taken way.”

In this way, Einstein dealt with the momentum of a quantum-theoretical radiation and the recoil of a molecule by the radiation, showing that the Maxwell distribution law of molecular velocity is thereby effected. As we have described in Sec. 3, Chap. 1 of the present volume, Einstein himself advocated in 1909 the necessity of recovering the wave nature of radiation, though he was of the standpoint of the light quantum theory. On these lines, in his paper in 1916 under consideration, he did not use explicitly the word light quantum. However, his expressions cited above, that “the construction of a theory of radiation having proper quantum nature appears almost inevitable” and that his theory “does not bring us nearer to the connection with the wave theory”, show that he aimed at going to a new theory of radiation based on the standpoint of the theory of light quanta.

Einstein introduced the concept of transition probability first for the out-radiation on the analogy of $\gamma$-radioactivity. $\gamma$-radioactivity was, similarly to $\alpha$- and $\beta$-radioactivities, an entirely new phenomenon, that was known to arise from the decay of individual atoms obeying a statistical law.\(^{43}\) From

\(^{43}\)The $\gamma$-ray from a radioactive element was found in 1906 [P. Villard, Comptes Randus 130 (1906), 1010, 1178]. It was considered to be a kind of electromagnetic wave similar to the X-ray. Its wavelength was measured for the first time in 1914 with the method of diffraction through crystal [E. Rutherford and E. N. da C. Andrade, Phil. Mag. 27 (1914), 854]. It was pointed out by Curie, Perrin, Becquerel, et al., that the constancy of atom would be violated if electrically charged $\beta$-rays were continuously emitted from an atom. From the measurement of the specific electric charge of $\alpha$-ray, it was found that the $\alpha$-particle was “a $\text{He}$ atom, more correctly, a $\text{He}$ atom when it loses the electric charge” [E. Rutherford
the standpoint of the theory of light quanta the emission of a γ-ray should be due to the decay of an individual atom accompanied by the emission of a light quantum. The transition between atomic states, of which W. Wilson gave the description "catastrophic" and Sommerfeld the description "of free will (freiwillig)", was also a newly found process accompanied by the emission of a light quantum. The transition probability, which Einstein introduced by saying that "one could hardly think of it other than according to the way of radioactive reaction", is a probability about a single atom like the decay probability, but is not a probability statistically averaged over many atoms.

From the viewpoint up to that time of what a theory should be, it was certainly a "weakness" of his theory that the time and direction of an individual elementary process was "entrusted to chance". This is because it was impossible to determine at what time and in what direction the elementary process would occur. Nevertheless, Einstein thought much of "a full trust in the reliability of the way" that he had taken. He did not restrict the way to overcome the "weakness". In other words, he said no more than that the transition probability "would be directly calculated", if we were "in possession of an electromagnetics and a mechanics modified in the meaning of the quantum hypothesis". He never said such a thing that the probability was only meaningful, or that it was an appearance coming from some inner-atomic cause. This was natural, because both the radioactivity and the transition between atomic states were quite newly known phenomena.

Einstein supposed that in the out-radiation the transition took place "without external influence (ohne äußere Einwirkung)". Bohr called this transition "spontaneous transition", in a 1918 paper which we shall treat in the next section. This word derives from the Latin word sponte (= of free will) and means having no external cause or influence, being synonymous with what Einstein meant. Sommerfeld called it "der freiwillig Übergang" (freiwillig = of free will).

Quoting Bohr's words "will start spontaneously to pass to the stationary state of smaller energy" in his paper just mentioned, Jammer writes in his book cited before in Sec. 4, §3, Chap. 2 of the present volume that in Bohr's words "spontaneously' is taken as essentially synonymous with 'acausal'", and that "in fact, following Bohr, Einstein's A_m" is often referred to as the 'probability factor of spontaneous emission', although Einstein never used this term.

and H. Geiger, Proc. Roy. Soc. London, 81 (1908), 162]. This made it clear that the exponentially decreasing radioactivity is due to the process, in which each individual atom decays with a certain probability by emitting the respective radioactive ray.
in his paper." After the discovery of quantum mechanics, there happened be­
tween Einstein and Bohr disputes on the interpretation of the wave function, in which a discussion was held on whether causality should exist in quantum mechanics or not. Jammer’s way of looking at the matter is, however, the one that speaks of the preceding event in terms of the later event of the so called Copenhagen interpretation of quantum mechanics, and makes a serious mis­take in the methodology of the history of science. What was made “in fact” by Bohr in the 1918 paper mentioned above was the proposal of the so called correspondence principle, in which the applicability of the classical electrodynamics to the calculation of $A_{m}^{n}$ for a transition between stationary states of large quantum numbers with a small difference was supposed.

4. The Correspondence Principle

§1. Angular momentum and the selection rule

As we have seen before in Sec. 2 of the present chapter, in those spectral lines that were derived from Sommerfeld’s quantum condition and Bohr’s frequency relation, there were included “extra lines” as Sommerfeld called them. Som­merfeld considered that in order to exclude them the combination rule should be applied conditionally. He put forth a “provisional” selection rule and Ep­stein modified it a little, but still no good enough agreement with experimental results was attained. Rubinowicz, at Sommerfeld’s institute, attempted to get more reasonable selection rules.

In his 1918 paper “Bohr’s Frequency Condition and the Conservation of Angular Momentum”,\(^{44}\) Rubinowicz calculates, with the use of classical electrodynamics, the energy and angular momentum of the radiation going out as a spherical wave from a moving electron, and takes into account the balance between this angular momentum and what the electron loses, arriving at the selection rule $|\Delta n_1| \leq 1$ for the azimuthal quantum number $n_1$. If the position vector of the electron is developed into a Fourier series of fundamental frequency $\nu$, the energy $\Delta W$ and the angular momentum $\Delta J$ going out in the form of spherical wave during one period $1/\nu$ are related by the equation

$$ |\Delta J| = \frac{\Delta W}{2\pi \nu} F, \quad (110) $$

\(^{44}\)A. Rubinowicz, Phys. Zeit. 19 (1918), 441, 465.
where $F$ is a numerical factor determined by the Fourier coefficients. When the electron makes a uniform circular motion, this reduces to $|\Delta J| = \Delta W/2\pi\nu$.\footnote{H. Butsch, Phys. Zeit. 14 (1913), 455; M. Abraham, Phys. Zeit. 15 (1914), 914.}

Rubinowicz next considers the transition of the electron from the state with the absolute value of angular momentum of $m\hbar/2\pi$ to the state with that of $n\hbar/2\pi$. Denoting the angle between the orbital planes before and after the transition by $\theta$, he has

$$|\Delta J| = \frac{h}{2\pi} \sqrt{m^2 + n^2 - 2mn \cos \theta}. \quad (111)$$

On the other hand, from Bohr’s frequency relation he has

$$\Delta W = h\nu. \quad (112)$$

Because Bohr’s frequency relation means that the radiation accompanying the transition is “strictly monochromatic”, the electron should make a “pure harmonic oscillation” according to the classical electrodynamics. This is in contradiction to the idea of the transition between two stationary states. On account of this, Rubinowicz assumes that “the radiating electron runs through an arbitrary periodic orbit whose size and form can change even slowly”.

Inserting Eqs. (111) and (112) into Eq. (110), and making use of the algebraic properties of $F$ given in terms of its Fourier coefficients, Rubinowicz shows that

$$|m - n| \leq 1. \quad (113)$$

He shows, furthermore, that the condition $m - n = \pm 1$ corresponds to the case of two circular orbits in a common plane, so that there result the selection rule $m_1 - n_1 = \pm 1$ of the azimuthal quantum number and the circularity of the polarization of radiation. It is shown also that, on the condition $m - n = 0$, the electron makes a simple harmonic oscillation if the two orbital planes are the same, giving rise to the selection rule $m_1 - n_1 = 0$ and linear polarization. These results are in agreement with the experimental facts on the Zeeman and Stark effects of which the theoretical investigations were made by Sommerfeld and Epstein.

Inserting $|\Delta J| = h/2\pi$, which comes from the quantum condition for the motion of an electron, into $|\Delta J| = \Delta W/2\pi\nu$ for the radiation from an electron moving in a circular orbit, Rubinowicz obtains the relation $\Delta W = h\nu,$
which shows the quantum nature of the energy radiated in the form of spher­
ical wave. That is to say, the quantum nature of the energy of light wave
is derived by the mediation of the quantum condition. In his paper in the
preceding year 1917, "On the Quantization of the Cavity Radiation", Rubinowicz has noticed Debye's derivation of the Planck distribution law, and
expressed the electromagnetic field in a cubic cavity as the superposition of normal oscillations of frequencies $\nu_i$'s. Taking the amplitudes $q_i$'s of the respective normal oscillations as the canonical coordinates to be used in the quantum conditions $\int p_i dq_i = n_i h$, etc., he has shown that the energy of the cavity radiation is given in the form $E = \sum n_i h\nu_i$. This acts as the germ of the quantization of fields which follows the discovery of quantum mechanics of particles.

To Sommerfeld in 1922, “this mediating standpoint seems satisfactory”, in view of the problem of “the propagation of emerging light wave based on
the wave theory and the Maxwell equation”, “in spite of the difficulty that the quantum-natured absorption of light and the photoelectric effect present”. To him in 1924, on the contrary, “the newest development, especially Compton's research, seems to exclude such a mediation”. As will be seen later in Chap. 5, however, Duane derives in 1923 the Bragg formula of reflection of X-ray by a crystal, from the standpoint of the theory of light quanta, by applying the quantum condition to the crystal lattice. It thus becomes clear that the quantum condition also mediates between the particle nature of light and the wave nature of light.

§2. The correspondence principle and selection rule

As is mentioned in the preceding subsection, Rubinowicz assumed that “the radiating electron runs through an arbitrary periodic orbit whose size and form can change even slowly”. But, he did not prescribe how “slowly” they should change. Bohr proposed in 1918, slightly before Rubinowicz, a hypothesis that classical electrodynamics might be applicable to a transition whose relevant quantum numbers were large in value but small in relative difference. Later

---

47) Cf. Sec. 1, §3, Chap. 1 of the present volume.
48) A. Sommerfeld, "Atombau und Spektrallinien" (Vieweg, Brauschweig, 1922; dritten Auflage).
49) A. Sommerfeld, loc. cit. (1924, vierten Auflage).
50) Sommerfeld means the discovery of the Compton effect in 1923.
in 1921 he named this hypothesis the correspondence principle. As we have seen before in Sec. 2, §7 of the present chapter, about the Zeeman effect Bohr supposed in 1914 that the frequency relation "is replaced by" a relation involving the strength of the magnetic field. Sommerfeld, in contrast to Bohr, imposed in 1916 some conditions, namely selection rules on the combination rule, as we have already seen, too. This time Bohr derived selection rules from the correspondence principle.

In his review article "On the Quantum Theory of Line-Spectra", 51) Bohr takes notice of Einstein's transition probability. Einstein derived the proportional relation \( \varepsilon_m - \varepsilon_n \propto \nu \) by making use of the transition probabilities and the Wien displacement law. If the universal proportional constant is put equal to the Planck constant, Bohr's frequency relation and the Planck distribution formula can be obtained. As we have described in Sec. 3, §2 of the present chapter, Einstein pointed out in 1917 that Bohr's frequency relation turned out to be secure by virtue of Sommerfeld's and Epstein's completion. Bohr next mentions that "by reversing the line of argument, Einstein's theory may be considered as a very direct support of this relation".

Of Einstein's transition probability of out-radiation \( A_{n''}^{n'} \), Bohr gives an objective description as follows — "In analogy with the fact that in the ordinary electrodynamics such a system will without external excitation emit a radiation of frequency \( \nu \), Einstein assumes in the first place that on the quantum theory there will be a certain probability \( A_{n''}^{n'} dt \) that a system in a stationary state of a greater energy characterized by the letter \( n' \), in the time interval \( dt \) will start spontaneously to pass to a stationary state of a smaller energy, characterized by the letter \( n'' \)." 52) Bohr's description of Einstein's transition probability of in-radiation \( B_{n''}^{n'} \) is similarly an objective one. On the basis facts, Bohr reports that "Einstein does not introduce any detailed assumption as to the values of these constants".

After doing so, for the purpose of determining the probability of a spontaneous transition, Bohr develops a way of thinking which he presented in his 1913 paper on the atomic model. If one considers a transition between a state corresponding to \( n = n' \) and a state corresponding to \( n = n'' \), and assumes that \( n' \) and \( n'' \) are large while \( n' - n'' \) is small compared with \( n' \) and \( n'' \),

---

52) This is the original sentence, from which Jammer quoted those words that we have discussed in the last part of Sec. 3, §4 of the present chapter.
one has\(^5\)

\[ \nu = (n' - n'')\omega. \] \hspace{1cm} (114)

In a stationary state of a periodic system, the displacement of a particle \(\xi\) in any given direction may always be expressed by means of a Fourier series of a sum of harmonic oscillations in the form

\[ \xi = \sum_{\tau=1}^{\infty} C_\tau \cos 2\pi(\tau\omega t + c_\tau). \] \hspace{1cm} (115)

Bohr says that 'in the ordinary theory of radiation we should therefore expect the system to emit a spectrum consisting of a series of lines of frequencies equal to \(\tau\omega\), but, as it is seen, this is just equal to the series of frequencies which we obtain by introducing different values of \(n' - n''\).

By the way, Bohr writes in the first section of the article under consideration the following — “If we consider a transition between two stationary states, it is obvious at once from the essential discontinuity, involved in the assumptions I and II\(^6\), that in general it is impossible even approximately to describe this phenomenon by means of ordinary mechanics or to calculate the frequency of the radiation absorbed or emitted by such a process by means of ordinary electrodynamics. On the other hand, from the fact that it has been possible by means of ordinary mechanics and electrodynamics to account for the phenomenon of temperature-radiation in the limiting region of slow vibrations, we may expect that any theory capable of describing this phenomenon in accordance with observations will form some sort of natural generalization of the ordinary theory of radiation.”

Now, in order to obtain the necessary connection with the ordinary theory of radiation in the limit of slow vibrations, Bohr considers that “we must further claim that a relation, as that just proved for the frequencies, will, in the limit of large \(n\), hold also for the intensities of the different lines in the spectrum”. The coefficient \(C_\tau\) in Eq. (115) determines, on the ordinary electrodynamics, the intensity of radiation. Thus, Bohr supposes that “we must expect that for large values of \(n\) these coefficients will on the quantum

\(^{53}\) This relation corresponds to Eq. (78) in Sec. 4, §3 of the last chapter.

\(^{54}\) These assumptions I and II correspond to the fundamental assumptions in the Bohr model of the atom (cf. Sec. 4, §2, p. 75, Chap. 2 of the present volume). But, Bohr here uses the terms “transition” and “unifrequent radiation” in I and II, instead of “passing” and “homogeneous radiation”, respectively, in the fundamental assumptions.
The Correspondence Principle

theory determine the probability of spontaneous transition. Also for small values of \( n \), Bohr says that “we may expect that the amplitude of the harmonic vibrations will in some way give a measure for the probability of a transition”.

This way of thinking is named the correspondence principle later in 1921 by Bohr in his lecture. Bohr derives the selection rules from the correspondence principle. That is, “if for all motions of a given system the coefficients \( C_\tau \) are zero for certain values of \( \tau \), we are led to expect that no transition will be possible. for which is \( n' - n'' \) equal to one of these values.” A linear harmonic oscillator makes the simplest example. In this case, since \( C_\tau = 0 \) for \( \tau \neq 1 \), it results that the frequency of any radiation emitted or absorbed by a linear harmonic oscillator is equal to a constant frequency.

For conditionally periodic motions of a system of multiple degrees of freedom Bohr takes a multiple Fourier series similar to Eq. (115). and about its coefficients considers in the same way as the above. For example, in the case of a conditionally periodic system with an axis of symmetry, the multiple Fourier series takes the form

\[
x = \sum D_{\tau_1,\tau_2} \cos 2\pi \left\{ (\tau_1 \omega_1 + \tau_2 \omega_2 + \omega_3) t + d_{\tau_1,\tau_2} \right\},
\]

\[
y = \pm \sum D_{\tau_1,\tau_2} \sin 2\pi \left\{ (\tau_1 \omega_1 + \tau_2 \omega_2 + \omega_3) t + d_{\tau_1,\tau_2} \right\},
\]

\[
z = \sum C_{\tau_1,\tau_2} \cos 2\pi \left\{ (\tau_1 \omega_1 + \tau_2 \omega_2) t + c_{\tau_1,\tau_2} \right\},
\]

where \( \omega_1, \omega_2 \) and \( \omega_3 \) are the three frequencies characterizing the system, and the \( z \)-axis is taken as the axis of symmetry. According to ordinary electrodynamics, the radiation from a charged particle moving in this way would therefore consist of a number of components of frequency \( |\tau_1 \omega_1 + \tau_2 \omega_2| \) polarized parallel to the axis of symmetry, and a number of components of frequencies \( |\tau_1 \omega_1 + \tau_2 \omega_2 + \tau_3 \omega_3| \) polarized circularly round the axis of symmetry when viewed in the direction of this axis.

Bohr then takes as the relation corresponding to Eq. (114)

\[
\nu = \sum \omega_k (n_k' - n_k''),
\]

on the assumption that

\[
n_k' \gg n_k' - n_k'',
\]

\((k = 1, 2 \text{ and } 3 \text{ in the above example})\). Bohr explains Eq. (119) in connection with the adiabatic hypothesis, as will be described in the next subsection.
§3. According to the correspondence principle, in the quantum theory the azimuthal quantum number $n_3$ would not change to emit a linearly polarized radiation, and would change by $\pm 1$ to emit a circularly polarized radiation.

In this way, with the use of the selection rules resulting from the correspondence principle, Bohr arrives, as in the case of Rubinowicz, at the removal of the lines "in discrepancy with the combination rule" as in Sommerfeld's words, and explain the semi-empirical rules of the polarization of radiation found by Epstein. Bohr also says that considerations according to the correspondence principle "allow also to throw light on the question of the relative intensities" of spectral lines. This task is accomplished in Kramers's paper, which we shall treat in §5 of the present section.

§3. The adiabatic hypothesis

In Sec. 1, §3 of the present chapter, we treated Ehrenfest's 1913 paper on the specific heat of the gas of diatomic molecules. In this paper, he showed that the rotational energy of a diatomic molecule should be an integer multiple of $\hbar \nu / 2$, $\nu$ being the number of rotations per unit time, by supposing that the rotational motion would be "adiabatically" realized in the limit of gradual weakening to nil of the external field to force the direction of the molecule oscillating sinusoidally.

In a 1914 paper "Contribution to Quantum Theory", Einstein called this idea of Ehrenfest "Ehrenfest's adiabatic hypothesis". In this paper Einstein discussed Nernst's theorem, that is, the third law of the thermodynamics. Calling some carrier of inner molecular energy "resonator" temporarily and assuming the energy of a resonator to be $\sigma \hbar \nu (\sigma = 0, 1, 2, \ldots)$, he showed first that the average thermal energy of a resonator is given by the Planck formula, from the fact that at the temperature $T$ the number $n_\sigma$ of molecules with energy $\sigma \hbar \nu$ is given in the form $n_0 e^{-\sigma \hbar \nu / kT}$. He next showed, assuming the molecule to take only discrete values $\varepsilon_\sigma$ of energy, that the entropy $S$ of a system of a given number of the molecules satisfies Boltzman's formula $S = k \log Z$, from the calculation of $S$ with the use of the number of the molecules $n_\sigma = n_0 e^{-(\varepsilon_\sigma - \varepsilon_0)/kT}$. In this calculation, $Z$ is defined to be "the number of the elementary states which are possible in the meaning of the quantum theory".

Einstein pointed out then that, if the thermodynamical state of the system depends on, besides the temperature $T$, a parameter $\lambda$ such as the volume, for example, $Z$ remains invariant for an adiabatic change of $\lambda$ from Ehrenfest’s adiabatic hypothesis, so that $S$ also remains invariant to ensure the Boltzmann principle in the quantum theory. Einstein expressed Ehrenfest’s adiabatic hypothesis as that “in an reversible adiabatic change of $\lambda$ every quantum-theoretically possible state pass again to such a state”. Einstein mentioned also that Ehrenfest’s adiabatic hypothesis is “a meaningful generalization of Wien’s displacement law”.

The adiabatic hypothetical way of thinking traces back originally to Einstein’s saying at the First Solvay Conference in 1911. In the discussion on Einstein’s lecture, Lorentz argued that when the string of a simple pendulum, which had at the beginning the energy quantum corresponding to its original period of oscillation, was shortened, its energy would become at last smaller than that energy quantum corresponding to its new frequency. Einstein’s answer to this was that “when the length of the pendulum changes endlessly slowly, the oscillatory energy remains equal to $h\nu$, if it has been equal to $h\nu$; the oscillatory energy changes like $\nu$”, and that “the same holds for a resistanceless electric oscillation and for free radiation”.$^{56}$ The case mentioned for free radiation means of course Wien’s displacement law. In this way, Einstein pointed out that, if the oscillatory energy is denoted by $E$, $E/\nu$ is adiabatically invariant.

Ehrenfest attempted, in his 1916 paper “Adiabatic Invariants and Quantum Theory”,$^{57}$ to understand Sommerfeld’s quantum condition, by thinking that in the adiabatic hypothesis much importance should be attached to adiabatic invariants in the classical mechanics. He took the case in which the potential energy of a system depended not only on the coordinates $q_1, q_2, \ldots$, but also on certain auxiliary variables $a_1, a_2, \ldots$, and the kinetic energy was a homogeneous bilinear function of $\dot{q}_1, \dot{q}_2, \ldots$. He called the process of changing the auxiliary variables infinitely slowly “reversible adiabatic change”. With the use of the Lagrange equation of motion, it was shown that for a system performing periodic motion, $\frac{2T}{\nu}$ is an “adiabatic invariant”, that is, it remains constant during the adiabatic change, $T$ and $\nu$ being, respectively, the average of the kinetic energy over time and the frequency.

$^{56}$ The Proceeding of the first Solvay Conference [the same as Ref. 11) on p. 46 in Chap. 2 of the present volume], S. 364.

Expressed in terms of the action integral, $\frac{2T}{\nu}$ is given by

$$\frac{2T}{\nu} = \sum \int p_k dq_k = \sum \int p_k dq_k.$$  \hspace{1cm} (120)

He supposed that "if the system of coordinates was well selected", the respective $\int p_k dq_k$'s in the expression for $\frac{2T}{\nu}$ would be adiabatic invariant.\(^{58}\)

Now, for a one-dimensional harmonic oscillator $2T$ is equal to the energy of the system $\varepsilon$, which is given by $\varepsilon/\nu = nh$ according to Planck's hypothesis of elementary quantum of energy. Ehrenfest then expected from the adiabatic hypothesis that for any periodical trajectory in the $(p_k q_k)$ plane, which could be brought about by means of a reversible adiabatic process from a harmonic motion, the quantum condition $\int p_k dq_k = nh$ would hold. We may note, however, that the quantum condition does not result only from the adiabatic hypothesis, just as Wien's distribution formula and hence Planck's distribution formula do not result only from the Wien displacement law. We must introduce the energy quantum, or the quantum condition for the harmonic oscillator. In this connection, we must not forget those respective substantialistic processes arriving at Wien's distribution formula and at Sommerfeld's quantum condition, of which we have already given a detailed description.

§4. The correspondence principle and adiabatic hypothesis

Now, returning to Bohr's review article, we see that he explains Eq. (119) given above in §2 as follows. Considering first a system with a single degree of freedom, Bohr calculates the variation $\delta I$ in the action integral $I$ when some new periodic motion is formed by a small variation in the first periodic motion. $I$ is given by

$$I = \int_0^\sigma p\dot{q} \, dt = 2 \int_0^\sigma T \, dt,$$  \hspace{1cm} (121)

where $T$ is the kinetic energy and $\sigma$ the period of the motion. For the new motion to be mechanically possible, there may be needed the existence of

\(^{58}\)It was verified in 1917 by J. M. Burgers [Versl. Akad. Amsterdam 25 (1917), 1055], that the individual action integrals are adiabatic invariants, for an adiabatic change in which the conditional periodicity of the system is conserved.
external force. By means of the Hamilton equation of motion, it is shown that

$$\delta I = \int_0^\sigma \delta Edt,$$

(122)

where $E$ denotes the energy of the system. On the assumption that the small variation in the motion is produced by a small and slowly varying external field, $\int_0^\sigma \delta Edt$ is shown to be zero, so that from Eq. (121) it results that $\delta(2T/\omega) = 0$ in terms of the frequency $\omega = 1/\sigma$, as in Ehrenfest’s work.

In the case of “systems for which every orbit is periodic independently of the initial conditions”, “we may for the varied motion take an undisturbed motion of the system corresponding to slightly different initial conditions”. This gives a constant $\delta E$, and one has from Eq. (122)

$$\delta E = \omega \delta I.$$

(123)

Bohr Eq. (114) in §2 on this equation. That is, if, in the transition from a state with quantum number $n'$ to that with quantum number $n''$, $n' - n''$ is small compared with $n'$ and $n''$, “we may for $\delta E$ put $E' - E''$ and for $\delta I$ put $I' - I''$, and get, with the use of Bohr’s frequency relation and the quantum condition $I = nh$,

$$\nu = \frac{1}{h} (E' - E'') = \frac{\omega}{h} (I' - I'') = (n' - n'') \omega.$$

Next, for a conditionally periodic system which makes motions of multiple periods, Bohr gives

$$\int_0^\theta \delta Edt = \sum N_k \delta I_k,$$

(124)

by making use of the Hamilton equations of motion. In this equation, $\theta$ is a sufficiently long time and the configuration of the system at $t = \theta$ is supposed, in a close approximation, to be the same as that at $t = 0$. $N_k$ is the number of oscillations performed by $q_k$ in the time interval $\theta$, and $I_k$ is the action integral $\int_0^{\sigma_k} p_k dq_k$ taken over the period $\sigma_k$ of $q_k$.

“In the special case where the varied motion is an undisturbed motion belonging to the same system as the unvaried motion, we get since $\delta E$ will be constant,”

$$\delta E = \sum \omega_k \delta I_k,$$

(125)
where $\omega_k = N_k/\theta$ is the mean frequency of oscillation of $q_k$. Bohr says that “this equation forms a simple generalization” of Eq. (123).\textsuperscript{59) For the case of a transition from a state with quantum numbers $n_k'$ to that with quantum numbers $n_k''$, where $n_k' - n_k''$ is small compared with $n_k'$ and $n_k''$, Bohr supposes again that we may put $E' - E''$ for $\delta E$ and $I_k' - I_k''$ for $\delta I_k$, and, make use of Eq. (125) resulting from classical dynamics, to get

$$\nu = \frac{1}{\hbar} (E' - E'') = \frac{1}{\hbar} \sum \omega_k (I_k' - I_k'') = (n_k' - n_k''),$$

by making use of the frequency relation and the quantum condition. This is Eq. (119) shown before in §2.

Bohr tries to understand Sommerfeld’s quantum condition by the adiabatic hypothesis. Bohr mentions that “this assumption is of great importance in the discussion of the conditions to be used to fix the stationary states of an atomic system among the continuous multitude of mechanically possible motions”. But, Bohr does not use the name of adiabatic hypothesis, but denotes it as “the principle of mechanical transformability”.

Bohr discusses also that if we consider at the beginning such a “system in which the forces which act on the particles are very small and where we may assume that the values of the energy in all the stationary states will tend to coincide”, “a continuous mechanical connection between two stationary states ... is just afforded by Ehrenfest’s principle”. About this point the reader may remember, however, that Ehrenfest’s adiabatic hypothesis supposes that every quantum-theoretically possible state passes through an adiabatic process into another such state, as Einstein pointed out, but does not assert the continuity among the states newly reached. We must say that the core of the correspondence principle consists in the thinking to see a kind of continuity for the transitions between stationary states which are essentially discontinuous, and to obtain the transition probability by means of classical dynamics.

Bohr’s thinking is based on the fact, shown in his 1913 paper on the atomic model, that the relation $\nu \simeq n \omega$ for the transition from an orbit with quantum number $N$ to one with quantum number $N - n (N \gg n)$ is the same as the relation between the frequency of the $n$th Fourier component of the orbital motion and that of the orbital motion itself.\textsuperscript{60) In Bohr’s treatment, Eq. (119) is

\textsuperscript{59) According to Schwarzschild's paper in 1916 which we have treated in Sec. 2, §6 of the present paper, we have $n_j = \partial F/\partial \alpha_j$ ($n$ is the frequency, $F$ the energy and $\alpha_j$ the action variable; cf. p. 124), so that we get at once $\delta F = \sum n_j \delta \alpha_j$, the same as Bohr's Eq. (125).

\textsuperscript{60) Cf. Sec. 4, §3, Chap. 2, p. 91 of the present volume.
grounded on Eq. (125). The latter equation has, however, no direct connection with the adiabatic hypothesis. It has only a certain similarity in the way of calculation to the theorems derived for the adiabatic invariants by Ehrenfest. Use is made of the concept of infinitesimal change in classical dynamics to express Bohr's supposition that, for transitions between two stationary states having quantum numbers large in values but small in difference, the results deduced from classical dynamics could be approximately applied. In this connection we may note that Bohr writes that “from the point of view taken in the present paper”, the name of the principle of the mechanical transformability “might in a more direct way indicate the content of the principle and the limits of its applicability”.

§5. The correspondence principle and intensities of spectral lines

Kramers takes up the task of calculating the intensities of the hydrogen lines in accordance with the correspondence principle, in a paper in 1919, “Intensities of Spectral Lines”. To express the conditionally periodic motion of the electron by the multiple Fourier series, Kramers applies the method of separation of variables to the Hamilton–Jacobi partial differential equation, as in the works of Epstein and Schwartzschild. With the action variables $I_k$ and angle variables $w_k = \omega_k t + \beta_k$, the quantum conditions and the frequencies of the electron are given by $I_k = n_k h$ and $\omega_k = \partial E/\partial I_k$, respectively, where $E$ is the energy of the electron and $k = 1, 2, \ldots s$.

Hamilton's characteristic function $W$ takes the form $W = \sum W_k(q_k; I_1, \ldots I_s)$ through separation of variables. Any function $f(q_1, \ldots q_s)$ can be expressed as

$$f(q_1, \ldots q_s) = \sum C_{\tau_1, \ldots \tau_s} e^{2\pi i (\tau_1 w_1 + \cdots \tau_s w_s)},$$

in terms of the multiple Fourier series. The Fourier coefficient $C_{\tau_1, \ldots \tau_s}$ is given by

$$C_{\tau_1, \ldots \tau_s} = \int_0^1 \cdots \int_0^1 f e^{-2\pi i (\tau_1 w_1 + \cdots \tau_s w_s)} dw_1 \cdots dw_s.$$
Transforming the integration variables using \( w_k = \partial W / \partial I_k \), one has

\[
C_{\tau_1, \ldots, \tau_s} = \int \cdots \int f \exp \left( -2\pi i \sum \sum \tau_k \frac{\partial W_j}{\partial I_k} \right) \Delta dq_1 \cdots dq_s , \tag{128}
\]

where \( \Delta \) denotes the functional determinant \( |\partial^2 W / \partial I_k \partial q_j| \).

Since it is necessary to treat the case of the transition from an initial state specified by a set of quantum numbers \( (n_1', \ldots, n_k'') \) to a final state specified by \( (n_1'', \ldots, n_s'') \), for which \( n_k' \) and \( n_k'' \) are large compared with \( (n_k' - n_k'') \), Kramers interprets Eq. (119), on which Bohr has grounded the correspondence principle, as follows. Kramers considers the multitude of states that satisfy the conditions \( I_k = \{ n_k'' + \lambda(n_k' - n_k'') \} \hbar, 0 \leq \lambda \leq 1 \). These states lie between the initial and final states. Those states for which \( \lambda \neq 0, 1 \) are not quantum-theoretical states, but are "mechanically possible states". The energy difference \( \delta E \) between the two mechanically possible states specified by \( \lambda \) and \( \lambda + d\lambda \) is given by \( \delta E = \sum \omega_k \delta I_k = \sum(n_k' - n_k'') \hbar \omega_k \lambda d\lambda \), because of \( \omega_k = \partial E / \partial I_k \). The frequency \( \nu \) of the radiation accompanying the transition is therefore given by

\[
\nu = \frac{1}{\hbar} \int_{\lambda=0}^{\lambda=1} \delta E = \sum(n_k' - n_k'') \bar{\omega}_k , \tag{129}
\]

where \( \bar{\omega}_k = \int_0^1 \omega_k d\lambda \) is the average value of \( \omega_k \) taken over all mechanically possible states.

When the two quantum numbers are large compared with their difference, one has \( \bar{\omega}_k = \omega_k \) because the motion in the initial state is almost the same as that in the final state. When the two quantum numbers are not large, the two states are different so that it becomes necessary, in view of Eq. (129), to take a certain mean value of the Fourier coefficients. Kramers says that the logarithmic mean value would be the natural one.

For the discussion of the Stark effect, he uses the parabolic coordinates \( (\xi, \eta, \phi) \) for \( (q_1, q_2, q_3) \) and calculates the Fourier coefficients of \( x + iy \) and \( z \), with the use of the relations \( x + iy = \sqrt{\xi e^{i\phi}} \) and \( z = (\xi - \eta)/2 \). The former contains only terms of the form \( e^{2\pi i(\tau_1 \omega_1 + \tau_2 \omega_2 + \tau_3 \omega_3)} \), and the latter only those of the form \( e^{2\pi i(\tau_1 \omega_1 + \tau_2 \omega_2)} \). From this, according to Bohr, Kramers derives the selection rules \( n_3' - n_3'' = \pm 1 \) for the "perpendicular components", polarized perpendicularly to the external electric field, and \( n_3' - n_3'' = 0 \) for the "parallel components", polarized in parallel to it.
Kramers names the ratio of the Fourier coefficient concerned to the major axis of the stationary orbit "the relative amplitude" and denotes it by $R$. If the number of the hydrogen atoms in the initial state of every line is assumed the same, $|R|^2$ gives an estimate of the relative intensities of the lines. He compares the theoretical relative intensities with the experimental ones of $H\alpha$, $H\beta$, $H\gamma$ and $H\delta$ of the Balmer series. In almost all the cases, the values of $R$ turn out to be zero. The lines split off by the Stark effect are indicated by $\Delta = n'(n_1' - n_2') - n''(n_1'' - n_2'')$, as was shown by Epstein and Schwartzschild [cf. Eq. (66) of the present chapter; $n = n_1 + n_2 + n_3$]. Kramers compares, as is shown in the figure, the arithmetic mean $(R'^2 + R''^2)/2$ of the values $R'^2$ and $R''^2$ of $R^2$ calculated for the initial and final states, respectively, with the experimental values.\(^{62}\)

For transitions of the type $(n_1', 0, n_3') \rightarrow (0, n_2'', n_3')$, it turns out that both $R'$ and $R''$ are zero. Nevertheless the perpendicular components of $H\beta$ with $\Delta = 10$, $H\gamma$ with $\Delta = 17$, and $H\delta$ with $\Delta = 26$ are of this type, they are observed experimentally though they are weak. Kramers argues that, since $R \neq 0$ for the mechanically possible states of $0 < \lambda < 1$, transitions of this type are not necessarily impossible, even if $R' + R'' = 0$, so that the theoretical values of relative intensities will be improved if suitable mean values of $R^2$ are taken. In the comparison with the experimental results, the stationary states with $n_3 = 0$, that is, states in which the electron has no angular momentum around the axis parallel to the external electric field, are omitted. Kramers points out that if such states were taken into account, the appearance of additional components of rather strong intensities should be expected.

Of the fine structure of the spectral lines of He, Kramers compares the intensities of the line of 3203 Å ($n = 5 \rightarrow n = 3$) and that of 3203 Å ($n = 4 \rightarrow n = 3$) with Paschen’s experimental results.\(^{63}\) Paschen’s experiment was done with the method of photographic plate for continuous discharge or spark discharge. Every line in the fine structure should therefore be considered to be slightly split into components on account of a weak electric field, so that the sum of $R^2$ over such components is to be compared with the experimental value. The motion of the electron parallel to the electric field consists of that with the frequency $|\tau_1 \omega_1 + \omega_2|$, which is present also in the case of no electric field, and those with the frequencies $|\tau_1 \omega_1|$ and $|\tau_1 \omega_1 + 2\omega_2|$, whose amplitudes are proportional to the strength of the electric field. The spectrum consists


\(^{63}\)F. Paschen, Ann. d. Phys. 50 (1916), 901.
therefore of the original components of \( n_2' - n_2'' = \pm 1 \) and the new ones of \( n_2' - n_2'' = 0, \pm 2 \). Kramers makes the comparison with the experimental values, noting that because the relative amplitudes of the "new" components are "of widely different orders of magnitude in the initial and in the final state", the sums of their \( R'^2 \) and \( R''^2 \) could not give more than an estimate of the order of magnitude of the intensities. Paschen's experimental results are of resolutions width roughly the same as the split width of spectral lines corresponding to an electric field of the inferred strength of 300 V/cm, but are in general accordance with the results of the calculation.

In this case, too, there are transitions for which both \( R' \) and \( R'' \) become zero, as in the case of the Stark effect. Kramers considers, however, if the mean value of \( R^2 \) over the mechanically possible states is taken for such a transition, the transition probability remains different from zero, and the difficulty mentioned above of the wide difference in the order of magnitude between the relative amplitudes in the initial and final state would be resolved. He points out the difficulty that in the case of the Zeeman effect of the line with the minimum quantum number, that is, the \((n = 2 \rightarrow n = 1)\) line of the Lyman series, the relative intensities of the split components remain mutually different, even if any mean value of \( R' \) and \( R'' \) is taken.

Kramers concludes that in spite of "the incomplete and preliminary character of underlying consideration", the results obtained on the Stark effect "must be considered as affording a general support of Bohr's fundamental hypothesis of the connection between the intensities of spectral lines and the amplitudes of the harmonic vibrations into which the motion of the electron in the atom may be resolved".

It should be noted, however, Kramers' analysis showed, at the same time, that the transition probabilities for essentially quantum-theoretical transitions related to small quantum numbers are not determined by \( C_{\tau_1,...,\tau_s} \), concerned with either one of the initial and final states, but should be prescribed in connection with both the initial and final states in some way, as the mean value considered by Kramers over the mechanically possible states. The transition probability \( A_{m}^{n} \) introduced by Einstein is of such a kind from the beginning. Here is seen the limitation of the correspondence principle in the sense that transition probability could be approximately given in terms of the Fourier coefficient \( C_{\tau_1,...,\tau_s} \). Later this problem becomes comes to be by means of the matrix element.
The Correspondence Principle

par. and perp. refer to the parallel and perpendicular components, respectively.

The numbers are values of $\Delta$. 
Chapter 4

Theory of Atomic Structure and Spin of Electron
Bohr (1913) Theory of atomic structure by means of the configuration in r(d) orbits.

Kossel (1916) Extension of the concept of the octet by arranging the electrons on concentric spherical shells.

Lewis (1916) Theory of chemical bond based on the model of cubic atoms (the octet theory).

Langmuir (1919) Extension of the octet theory by arranging the electrons on concentric spherical shells.

Wentzel (1921) Distinction between screened doublet terms and relativistic doublet terms with two kinds of quantum numbers.

Coster (1921) The same as the above by the kinds a, b.

Herz (1920) Discovery of screened doublet terms of X-rays.

Sommerfeld (1921) Similarity of the quantum number for doublet terms of X-rays.

Wentzel (1921) Similarity of the quantum number for X-ray of a element to the quantum number for X-ray of the preceding element.

Sommerfeld (1920) Introduction of inner quantum number for doublet and triplet terms.

Sommerfeld (1916) Similarity of alkali doublet terms to relativistic doublet terms.

Kossel (1919) Confirmation of At.

Barnett (1915), de Haas (1915) Measurement of magnetic effect due to rotation and the angular momentum of a ferromagnetic body.

van der Pauw (1948) Corning of a magnetic moment due to rotation and angular momentum of a ferromagnetic body.

Landé (1920) Analysis of anomalous Zeeman effect with half-integral quantum numbers.

Zener (1922) Similarity of the spectrum of an element to that of the preceding element.

Sommerfeld (1922) General spectrum formula.

Schrödinger (1926) Theory of chemical bond by means of circular orbits containing up to eight electrons.

Heisenberg (1922) Division of orbital angular momentum into half-integral quantum numbers.

Coster (1922) Discovery of multiplet terms.

Stark (1922) Confirmation of atomic magnetic moment. M"{o}gner of a magnetic moment splitting.

Becquerel (1921) Systematic analysis of anomalous Zeeman effect by means of quantum numbers.

Larmor (1917) Analytical of anomalous Zeeman effect of the type of an element. Larmor frequency.

Runge (1907) Of the anomalous Zeeman effect = a fraction of the line frequency.
Bohr-Coster (1923)  
Introduction of two kinds of azimuthal quantum numbers \( k, \bar{k} \).

Bohr (1923)  
Introduction of nonmechanical compulsion to cause alkali doublets.

Millikan-Bowen (1924)  
Confirmation of the equality pointed out by Sommerfeld. Indication of the contradiction between \( n(k, \bar{k}) \) configuration and relativistic fine structure.

Landé (1923)  
Study of the equality of the two kinds of terms. \( k_\text{r} \) = inner quantum number.

Landé (1923)  
Composition of the angular momenta of an electron and the atomic body \( \rightarrow \) g-factor formula.

Millikan-Bowen (1924)  
Confirmation of the above equality. Interpretation of relativistic doublet terms by magnetic force between electron and atomic body.

Störmer (1924)  
\( n(k, j) \) electron configuration.

Pauli (1925)  
Relativistic effect on g-factor. Doubleness of electron to cause alkali doublet terms.

Uhlenbeck-Goudsmit (1925)  
Introduction of electron spin. Interpretation of doublet terms by spin-orbital interaction.

Sucksmith-Bates (1923)  
Confirmation of magneto-mechanical anomaly \( \rightarrow \) twice the value for \( e/m \).

Notes:  
1) The numbers above a frame indicate Sec. and Subsec. 3-means Chap. 3  
2) \( n \): principal quant. number,  
\( k \): azimuthal quant. number,  
\( j \): internal quant. number,  
\( e/m \): electron's specific charge.
Chapter 4

Theory of Atomic Structure and Spin of Electron

Theme in Chapter 4

One of the main aims of the quantum theory was to clarify the atomic structure by making use of the various atomic phenomena as clues. On the basis of the success of Bohr's theory of the hydrogen atom, many kinds of challenge were made one after another. Generally speaking, the clarification of the atomic structure was proceeded through two ways. One way was based on the results in the spectroscopy of X-rays as well as light. The other was one that brought the phenomenological summaries of valuable chemical knowledge, such as represented by Mendelev’s table of periodicity in 1869, to a substantialistic stage by analyzing them in connection with the atomic structure. This was the stage that became at length realizable on account of the progress in the quantum theory.

The old quantum mechanics, which was developed on the basis of the quantum condition and selection rule, encountered with the difficulty of the factor 2. That is, firstly, the analyses of the anomalous Zeeman effect required half-integer quantum numbers and the anomalous coefficient 2 of magnetic moment. Secondly, in Stern-Gerlach’s experiment, which was interpreted as proving the directional quantization, the atomic beam split into two components, instead of three components as in the normal Zeeman effect. Thirdly, the results of experiments on magnetization of matter demanded such a value for the specific charge of the electron as was twice as large as the value hitherto known.

The analysis of X-ray spectra brought into existence another kind of spectral fine structure different from the relativistic fine structure. Of the substantialistic processes in which these problems were thrown into relief, not enough analyses have hitherto been made in the field of the history of science. Of the developments through which these problems were brought to solution, more or less similar situation has been seen.

Bohr was in doubt about the half-integer quantum number and thus sought to solve the problem in terms of “non-mechanical constraint”, introducing two kinds of azimuthal quantum numbers to treat the atomic structure. It was made clear by Landé and by Millikan and Bowen, however, that this was contradictory to experimental observations. Stoner then revised, with the use of half-integer quantum number, Bohr's scheme of configuration of electrons in atoms. This revision opened the way to Pauli's proposal of
The Anomalous Zeeman Effect and Half-Integer Quantum Numbers

1. The Anomalous Zeeman Effect and Half-Integer Quantum Numbers

§1. Inner quantum numbers and half-integer quantum numbers

§1.1 Doublet and triplet terms and inner quantum numbers

Having succeeded in explaining the spectral series of hydrogen using the quantum condition, Sommerfeld made a number of efforts to understand spectral series of other atoms in terms of the quantum condition.

In his 1916 paper "On the Quantum Theory of Spectral Lines; Supplement and Extension", Sommerfeld considers the energy $E$ of an atomic electron by means of a model in which an electron in a neutral atom is assumed to move outside the group of the other electrons. He calculates the electrostatic energy of the electron concerned, by supposing that the group of the other electrons are distributed homogeneously along a circle of radius $a$ around the atomic nucleus as center. Expanding the electrostatic energy as a power series in $a$, he then calculates the action integral $\int p_r \, dr$ to get, in the approximation up to $a^4$,

$$E = -\frac{Nh}{\{k + n_r + q_k + \kappa_k (k + n_r q/k^3)^{-2}\}^2}.$$  \hspace{1cm} (1)

In this equation, $k$ is the azimuthal quantum number, $n_r$ is the radial quantum number, $q$ is a constant determined by $a$ and the amount of electric charge on the circle, and $q_k$ and $\kappa_k$ are quantities dependent on $k$ and $q$. For large enough $k$, one has $q_k \simeq q/k^3$ and $\kappa_k \simeq 0$. $N$ is the Rydberg constant.

This result is in good accordance with Ritz’s empirical law of spectral terms, that is, Eq. (52), Chap. 2 of the present volume. We have already mentioned that the spectral term $(m, \alpha, \beta)$ given by Eq. (52), Chap. 2 is called the term $ms$ for sharp series, the term $mp$ for principal series, and the term $md$ for diffuse series. The term $mf$ was further known for fundamental series.

---

(Bergmann series). In consideration of $k \geq 1$ and $n_r \geq 0$, the frequency $\nu$ of a spectral line in these series is expressed with $n = k + n_r$ as

- **Principal series**  $\nu = 1s - np, \quad n = 2, 3, 4, \ldots$
- **Sharp series**  $\nu = 2p - ns, \quad n = 2, 3, 4, \ldots$
- **Diffuse series**  $\nu = 2p - nd, \quad n = 3, 4, 5, \ldots$
- **Fundamental series**  $\nu = 3d - nf, \quad n = 4, 5, 6, \ldots$

and the terms $s, p, d$ and $f$ correspond to $k = 1, 2, 3$ and $4$, respectively.

As Kossel and Sommerfeld point out in their paper “Selection Principle and Displacement Law in Series Spectra”,\(^2\) these spectral series correspond to the transitions $k \rightarrow k \pm 1$, being described according to the selection rule $k \rightarrow k, k \pm 1$. They confirm the fact that the spectrum of an element obtained by means of spark discharge is of the same structure as that of the preceding element in the periodicity table obtained by means of arc discharge, and call it the “displacement law in series spectra” (Verschiebungssatz bei Serienspektren). Sommerfeld\(^3\) says that “the meaning of our law is clear according to the model”. Because the elements in one and the same column of the periodicity table have an identical number of valence electrons, that is, an identical number of outermost electrons in the model, it suffices to consider that the spectrum of an ionized atom is seen in a spark discharge, while that of a neutral atom is seen in an arc discharge.

Corresponding to the doublet and triplet lines in spectral series, doublet and triplet spectral terms should also be taken into consideration. In the paper “General Spectroscopic Laws, especially a Magneto-optical Splitting Law”,\(^4\) Sommerfeld attaches importance to the fact that when a $p$-term is a doublet (or triplet) the $d$-term combined with it is also a doublet (or triplet), and called this the “permanence of multiplicity”. $s$-terms are always singlet. However, all the lines that are to arise from the combinations between doublet terms or between triplet terms, are not always observed. For example, the combination $\nu = 2p_i - 3d_j$ between the doublet terms of Cs gives, instead of four lines, two strong “lines” and a “satellite”. The combination $\nu = 2p_i - 3d_j$ between triplet terms of Cd gives six strong and weak lines, instead of nine lines.

---

\(^3\)A. Sommerfeld, Ref. 4) given below.
\(^4\)A. Sommerfeld, Ann. d. Phys. 63 (1920), 221.
Sommerfeld attempts in this paper to understand this fact by “a kind of selection principle”. The azimuthal quantum number \( k \) is to prescribe a term, not to select the individual \( p \)- or \( d \)-terms. Sommerfeld says that “the azimuthal quantum number is related to the angular momentum of the whole atom, that is to say, its external rotation”, and suggests that “the discriminating mark of various \( p \)- and \( q \)-terms must rather be an inner quantum number, something corresponding to a hidden rotation”.

From such a substantialistic consideration, Sommerfeld assigns as the internal quantum number \( j \), 1 for \( s \)-terms, 2 and 1 for doublet \( p \)-terms, 2, 1 and 0 for triplet \( p \)-terms, 3 and 2 for doublet \( d \)-terms, and 3, 2 and 1 for triplet \( d \)-terms, and supposes that there holds the same selection rule \( j \to j, j \pm 1 \) as that for the angular momentum. As far as this is concerned, these assigned values of \( j \) have the arbitrariness of an additive constant.

§1.2 *Anomalous Zeeman effect of doublet and triplet terms and half-integer quantum numbers*

Sommerfeld further notices Runge’s rule. Runge\(^5\) arranged observational results of the Zeeman effect, noting that, though “the normal splitting” into three components in a magnetic field can be explained by the classical theory,\(^6\) “for the anomalous splitting one has no formula”, and found that the separation \( \Delta \nu \) in frequency in a magnetic field of strength \( H \) “shows the following property”, namely

\[
\Delta \nu = \frac{(q/r)}{\Delta \nu_0}; \quad q = 1, \pm 1, \pm 2, \ldots \tag{2}
\]

In this formula, \( \Delta \nu_0 = eH/4\pi mc \) is the Larmor frequency which gives the separation in frequency in the Zeeman effect, and \( r \) is an integer prescribed to each original line. Sommerfeld assigns “Runge’s denominator” \( r \) in Runge’s rule of the anomalous Zeeman effect to the \( s \)-, \( p \)-, \( d \)-, \( f \)- and following \( x \)- and \( y \)-terms in accordance with the combination rule. That is, he puts \( r = 1 \) for singlet terms, \( r = k \) for triplet terms and \( r = 2k - 1 \) for doublet terms, and calls this a “mystery of number”.

Through an arrangement on the basis of Runge’s rule, of the observational results on the anomalous Zeeman effect in weak magnetic field, Back\(^7\) arrives

---

\(^5\)A. Runge, Phys. Zeit. 8 (1907), 232.
\(^6\)Cf. Sec. 2, §3, Chap. 1, Vol. I.
\(^7\)E. Back, Die Naturwiss. 9 (1921), 199.
in 1921 at a phenomenological lawfulness by which new lines may be predicted. Landé attempts, in his paper in “On the anomalous Zeeman Effect”, to grasp Back’s result with the use of Sommerfeld’s inner quantum number in accordance with the selection rule, and puts the change $\Delta E$ in energy of a spectral term due to a magnetic field as

$$\Delta E = mgh\Delta\nu_0,$$

(3)

where $m$ is given by in terms of the inner quantum number $j$ as

$$m = \begin{cases} 0, \pm 1, \pm 2, \ldots, \pm j & \text{(singlet and triplet terms),} \\ \pm \frac{1}{2}, \pm \frac{3}{2}, \ldots, \pm \left( j - \frac{1}{2} \right) & \text{(doublet terms),} \end{cases}$$

and $g$ is a fraction dependent on $j$ and the azimuthal quantum number $k$. $g$ is what later becomes known as Landé’s $g$-factor. $g$ takes the value $g = 2j/(2k-1)$, for example, for doublet terms ($j = k$ and $k-1$). Landé calls $m$ the “equatorial quantum number”. $k$ and $j$ are supposed to obey the selection rules

$$k - k' = \pm 1; \quad j - j' = 0, \pm 1,$$

except that $j = j' = 0$, and $m - m' = 0$ in the case of $k - k' = 0$ are forbid-den. Landé notices that the theory of the normal Zeeman effect giving $g = 1$ is based on the directional quantization, in which the angle $\Theta$ between the rotational axis of the atom and the direction of the external magnetic field is supposed to take only values given by $k \cos \Theta = m$ (integer) and $|m| \leq k$. He then “identifies the hypothetical ‘inner’ quantum number as the azimuthal quantum number of the atom as a whole in its invariable plane, contrary to the ‘outer’ azimuthal quantum number of a springing outer electron”, and supposes that $|m| \leq j$ if the angular momentum of the atom denoted by $j$ has the projection $m$ in the direction of the magnetic field. On account of Runge’s rule, “one must be satisfied with rational fractions for $m$”. Because $m$ changes by one according to the selection rule and takes positive and negative values symmetrically, Landé supposes that the series of values $|m| = 1/2, 3/2, \ldots |j - 1/2|$ comes into consideration, other than the series of values $|m| = 0, 1, 2, \ldots j$, and thereby introduces the half-integer quantum number.

By the way, as we shall describe in the present chapter and in Chap. 3, Vol. III, the half-integer quantum number introduced by Landé becomes later

8) A. Landé, Zeit. f. Phys. 5 (1921), 231, 7 (1921), 398; Phys. Zeit. 22 (1921), 417.
fully understood in terms of the spin of the electron, through hard efforts made by many physicists to propose various ideas. It is the main purpose of the present chapter to reveal the drama, so to speak, of such a process in the old quantum mechanics. The introduction of half-integer quantum numbers should be said to be the first grasp of the indication of the spin of the electron.

Now, Landé assigns the value 2 for his $g$-factor of the $s$-term. Landé points out that this is an "utterly incomprehensible result", because this gives "an apparent doubling of $e/m$", the specific charge of the electron. As we shall see in §3 below, experiments on ferromagnetic substances have already shown that the specific charge of the electron should be nearly doubled for the experimental results to agree with theory. Landé calls our attention to the connection of this fact with the anomalous value of the specific charge of the electron demanded by the $g$-factor.

Sommerfeld, in his 1922 paper "Quantum Theoretical Reinterpretation of Voigt's Theory of the Anomalous Zeeman Effect of D-line Type",\textsuperscript{9} treats quantum-theoretically Voigt's work\textsuperscript{10} on the Paschen-Back effect\textsuperscript{11} of the absorption spectrum of the doublet D-lines of Na, and assumes the following formula for the doublet term in general:

\[ \Delta E = \frac{1}{2} \left( m \pm \sqrt{1 + \frac{2m}{2k-1} u + u^2} \right) h\Delta \nu_0, \]  

(4)

where $u = \Delta \nu_D / \Delta \nu_0$, $\Delta \nu_D$ being the difference in frequency between the D\textsubscript{1}-, D\textsubscript{2}-lines. Because of the ratio 2 of the intensity of the D\textsubscript{2} line to that of the D\textsubscript{1}-line, Voigt considered a system of three electrons, two of them making quasi-harmonic oscillations of frequency $\nu_2$ and the other one that of frequency $\nu_1$. Each electron was assumed to be acted on, not only by a force proportional to the product of $eH/c$ with its velocity, but also by forces proportional to similar forces acted on the other electrons. From the classical-theoretical results on the differences between the eigenfrequencies of the system, Sommerfeld derives, by making use of the frequency-energy relation, a formula which is the same as Eq. (4) with $m = \pm 1, \pm 3$ and $2k - 1 = 3$. Because this value 3 is equal to Runge's dominator of the $p$-term giving the D-lines, Sommerfeld generalizes such a formula to obtain Eq. (4).

\textsuperscript{9}A. Sommerfeld, Zeit. f. Phys. 8 (1922), 257.
\textsuperscript{10}W. Voigt, Ann. d. Phys. 41 (1913), 403; 42 (1913), 210.
\textsuperscript{11}F. Paschen und E. Back, A. d. Phys. 39 (1912), 897.
Heisenberg points out in his 1922 paper “On the Quantum Theory of the Line Structure and the Anomalous Zeeman Effect”,\(^\text{12}\) that, if one puts \(m^* = m/2\), \(m^*\) takes the value \(\pm(1/2, 3/2, \ldots)\) and Eq. (4) becomes identical with Landé’s result in the case of weak magnetic field \((u \sim \infty)\), and becomes \(\Delta E = (m^* \pm 1/2)\hbar\Delta \nu_0\) in the case of strong magnetic field \((u \sim 0)\). Because \(m^* \pm 1/2\) takes integer numbers, this is in accordance with the fact found in the Paschen–Back effect.

§2. Half-integer quantum numbers and anomalous coefficients of magnetic moment

§2.1 Half-integer quantum number and “atomic body”

In order to explain the half-integer quantum number introduced by Landé, Heisenberg proposes, in the paper cited just above, a substantialistic model in which the outermost electron shares its angular momentum with the “atomic body” \((\text{Atomrumpf})\). He supposes that every atom of the alkalis, which has only doublet terms, consists of an outermost electron, that is a valence electron, and a group of other electrons tightly bounded around the nucleus, that is the atomic body. To the normal state of an alkaline atom an s-term is assigned, and the outermost electron has an angular momentum of value 1 in unit of \(\hbar/2\pi\). Heisenberg considers that this angular momentum is, in the course of the motion of the valence electron, exchanged between it and the atomic body. In an excited state of the atom to which the azimuthal quantum number \(k\) is assigned, the valence electron and the atomic body are assumed to have angular momenta \(k^* = k - 1/2\) and 1/2 respectively.

These two angular momenta combine to result in the total angular momentum \(j\) of the atom, and \(k^*\) makes a precession around \(j\). If an external magnetic field \(H\) is applied, the angle \(\Theta\) which \(j\) makes relative to \(H\) is determined by the component \(m^*_i\) of \(k^*\) in the direction of \(H\), and by the average value of \(k^*\) in the direction of \(j\) during the precession. The magnitude of \(j\) depends on the component \(m_i^*\) of \(k^*\) in the direction of the angular momentum 1/2 of the atomic body. The change \(\Delta E\) in energy of the atom due to the magnetic field \(H\) is given by \(h\Delta \nu_0 j \cos \Theta\). Heisenberg assumes that when \(m_i^*\) can take \(\lambda\) different values the sum of the values of \(\Delta E\) corresponding to each value of \(m_i^*\) is equal to \(\lambda h\Delta \nu_0 m^*\). He argues that, because the electromagnetic field carries

away an angular momentum of unit magnitude in a transition $\Delta m^* = 1$, the electromagnetic field may be supposed to carry away an angular momentum of magnitude $\lambda$, if all the $\lambda$ atoms having the various values of $m_i^*$ are taken into consideration.

From this condition, it is shown that only the two values $m_i^* = \pm k^*$ are allowed, and therefore the angular momentum of the valence electron is in a direction parallel or opposite to the angular momentum of the atomic body. It follows that $j = k^* \pm 1/2$, namely, $j = k$ or $k - 1$, which is identical with the inner quantum number of the doublet term. Heisenberg also evaluates the interaction energy between the magnetic field generated by the orbital motion of the valence electron and the magnetic moment of the atomic body, by applying the quantum condition to the orbital motion, and shows that it can explain the magnitude of the splitting of the doublet term.

Since the triplet term is seen in the case of an element that has two valence electrons, Heisenberg considers two possibilities for the share of angular momentum among the valence electrons and atomic body. One possibility is that one valence electron takes an angular momentum of $1/2$, the other electron takes $-1/2$ and the atomic body takes $0$. He remarks that this case corresponds to $1S$ state in spectroscopy. The other possibility is that every valence electron takes an angular momentum of $1/2$ and the atomic body $1$. In an excited state, the angular momenta of the inner valence electron and the atomic body are supposed to compose the angular momentum $3/2$, while the outer valence electron has $k^* = k - 1/2$. From a calculation of the change in energy of the atom when an external magnetic field is applied, done in a way similar to that used in the case of the doublet term, he concludes that $m^*$ could take only three values $0$ and $\pm k^*$.

Heisenberg attempts, in short, to explain the half-integer quantum number for the doublet term by supposing that the relation which Rubinowicz showed, as we have seen in Sec. 4, §1 of the preceding chapter, about an elementary process of transition of an individual atom might be applied on the average to the transitions of $\lambda$ atoms. Though this attempt does not prove to be appropriate, the idea that an atomic body or valence electron has a half-integer value of angular momentum is taken over by Landé and Sommerfeld.

---

13) In spectroscopy, the notations $S, P, D, \ldots$ are sometimes used for singlet terms, and $s, p, d, \ldots$ for doublet and triplet terms.
Sommerfeld and Heisenberg consider, in their paper "The Intensity of Multiple Lines and their Zeeman Components", the selection rule for the inner quantum number and the intensities of spectral lines in the Zeeman effect, on the basis of the correspondence principle. In this paper, \( j \) is considered to be the quantum number of the total angular momentum of an atom in a stationary state, after the results of Landé and Heisenberg. Sommerfeld and Heisenberg use the Fourier series expansion of the orbital motion of an electron in a plane designated by the coordinates \((\xi, \eta)\), and make a transformation into the coordinates \((x, y, z)\) whose \(z\)-axis is taken in the direction of \( j \). On the assumption that the orbital plane makes a precession around the direction of \( j \), they multiply \( x \) and \( y \) by the factor \( e^{i\omega_j t} \), \( \omega_j \) being the angular frequency of the precession. From the components of frequency thus obtained of \( x \), \( y \) and \( z \), they get the selection rule \( \Delta j = 0, \pm 1 \) according to the correspondence principle. They say that the rule \( j = 0 \rightarrow j = 0 \) is "doubtless", because \( j \) represents the total angular momentum.

About the "magnetic quantum number" \( m \), which is the component of \( j \) in the direction of the external magnetic field, they get the selection rule \( \Delta m = 0, \pm 1 \), by applying the transformation into the coordinates \((X, Y, Z)\) with the \(Z\)-axis taken in the direction of the magnetic field. They calculate the relative intensities of the component lines, and conclude that "the correspondence principle is verified also in its application to the proper Zeeman splitting."

§2.2 The generalization of \( g \)-factor and the anomalous coefficient of magnetic moment

In the same year 1922, Catalán\textsuperscript{15}) found in the spectrum of Mg a group of lines lying regularly, and called them "multiplet". They are, for example, nine lines that seem to result from the combination of a triple \( p \)-term with a quintet \( d \)-term, and fifteen lines that seem to result from the combination between a quintet \( d \)-term and a \( f \)-term. This shows that Sommerfeld's "permanence of multiplicity" does not hold as it is. Catalán finds multiplets also in the spectra of Ca, Sr, Ba, Al\textsuperscript{+}, Sc\textsuperscript{+}, Y\textsuperscript{+}, Si and Cr.

Landé denotes, in his paper "Term Structure and Zeeman Effect of Multiplets",\textsuperscript{16}) a multiplet term with the notation \( n_{kj}^r \), where \( n \), \( k \) and \( j \) are the

\textsuperscript{14)}A. Sommerfeld und W. Heisenberg, Zeit. f. Phys. 11 (1922), 131.
\textsuperscript{15)}M. A. Catalán, Phil. Trans. London (A) 223 (1922), 127.
\textsuperscript{16)}A. Landé, Zeit. f. Phys. 15 (1923), 189.
principal, azimuthal and inner quantum numbers respectively and \( r \) represents the multiplicity, being equal to 1, 2, 3, 4, \ldots for singlet, doublet, triplet, quartet, \ldots terms respectively. \( k \) obeys the selection rule \( k \rightarrow k \pm 1 \), and \( j \) the selection rule \( j \rightarrow j \pm 1 (j = 0 \rightarrow j = 0) \). Landé assumes the selection rule \( r \rightarrow r, r \pm 2 \). This means that Landé modifies Sommerfeld's "permanence of multiplicity". Introducing \( R = r/2 \) instead of \( r \), \( K = k - 1/2 \) instead of \( k \), and \( J = j \) for even multiplet terms and \( J = j - 1/2 \) for odd multiplet terms instead of \( j \), the selection rules transform into \( K \rightarrow K \pm 1, R \rightarrow R, R \pm 1 \) and \( J \rightarrow J, J \pm 1 \), all being of the same form. Landé supposes that \( J \) obeys the condition

\[
|K - R| + \frac{1}{2} \leq J \leq |K + R| - \frac{1}{2}.
\]

Thus, he has, for example, \( J = K \pm 1/2 \) for doublet terms for which \( R = 1 \).

Landé considers that \( K \) is the angular momentum of the radiating electron and \( R \) that of "the system of the body electrons". Their vector sum results in angular momenta of magnitudes

\[
R + K, R + K - 1, R + K - 2, \ldots, |R - K| + 1, |r - K|,
\]
on the assumption that it takes values differing stepwise by one. The values that are obtained by taking the arithmetical means of the adjacent pairs of these values coincide with the values of \( J \) given by Eq. (5). According to directional quantization the component \( m \) of \( J \) in the direction of the magnetic field has the values,

\[
J, J - 1, J - 2, \ldots, -(J - 1), -J.
\]

Taking the arithmetical means of the adjacent pairs of these values also, Landé assumes that \( m \) takes the values

\[
J - 1/2, J - 3/2, \ldots, -(J - 3/2), -(J - 1/2).
\]

Therefore, the number of values that \( m \) can take becomes \( 2J \) for a doublet term, for example, that is \( 2k \) and \( 2(k - 1) \), being \( 2(2k - 1) \) in total.

Landé puts \( g = (J + R \cos \angle RJ)/J \) in consideration of the precession of \( R \) around \( J \). Using \( \cos \angle RJ = (J^2 + R^2 - K^2)/2JR \) and replacing \( J^2 \) with \( \bar{J}^2 = (J + 1/2)(J - 1/2) \), he gets

\[
g = 1 + \frac{(\bar{J}^2 + R^2 - K^2)}{2\bar{J}^2}.
\]
\( \tilde{J} \) corresponds to the geometrical mean of those values whose arithmetical mean gives \( J \). With this choice of \( \tilde{J} \), the selection rule \( J = 1/2 \rightarrow J = 1/2 \) changes into \( \tilde{J} = 0 \rightarrow \tilde{J} = 0 \), which is in accordance with Sommerfeld-Heisenberg's selection rule for the inner quantum number.

By making use of \( g \) given above, the splitting \( mg\Delta \nu_0 \) of a term in the Zeeman effect agrees well with the respective experimental result. Though Landé considers that the angular momenta represented by half-integer quantum numbers “describe apparent angular momenta, from which the real mechanical momenta may differ”, he says that in the formulation \( g = (\tilde{J} + R \cos \angle R J)/\tilde{J} \), or Eq. (6), “a clear hint for the decisive solution of the problem of the anomalous Zeeman splitting seems to lie”.

About the inner quantum number \( j \), Sommerfeld in his 1923 paper “On the Interpretation of Complicated Spectra (Magnesium, Chromium, etc.) by the Method of the Inner Quantum Number”, supposes that it describes the total angular momentum of an atom in an excited state, which results from the vector sum of the total angular momentum \( j_s \) of the atom in the unexcited state and the angular momentum \( j_a \) of the excitation. \( j_a \) is supposed to be 0, 1, 2, \ldots for the \( s, p, d, \ldots \) terms respectively. That is, he puts \( j_a = k - 1 \). On the assumption that the relative direction between \( j_s \) and \( j_a \) is determined “in a way similar to the directional quantization in an external field”, the number of the values that \( j \) can take is \( 2j_a + 1 (j_a < j_s) \) or \( 2j_s + 1 (j_a > j_s) \). The number of levels that a spectral term can take increases, therefore, with \( k = j_a + 1 \), but remains constant when \( j_a > j_s \). Sommerfeld names this constant value \( 2j_s + 1 \) a “permanent number”.

In the case of alkaline elements, every atom in its normal state corresponds to an \( s \)-term, giving \( j_s = 1 \). The formulation mentioned above does not, therefore, explain the alkaline doublet terms. Sommerfeld thus supposes, following Heisenberg, that \( j_s + 1 \) is equally divided between the valence electron and the atomic body. For excited states, Sommerfeld considers the vector sum of \( j_a + 1/2 \) and \( 1/2 \). Its maximum and minimum values are \( j_a + 1 \) and \( j_a \) respectively, and there is “no third integer level lying between them”. He considers that in this way the doublet term is obtained, for which the inner quantum number takes the values \( j = j_a + 1 = k \) and \( j = j_a = k - 1 \).

Sommerfeld, in his 1924 paper “On the Theory of the Multiplets and their Zeeman Effects”, modifies this formulation, taking half-integer \( j_s \) for even

17) A. Sommerfeld, Ann. d. Phys. 70 (1923), 32.
terms. In this paper he calls \( j_s \) "the inner quantum number in an unexcited state". For an even term, a half-integer \( j_s \) and an integer \( j_a = k - 1 \) are supposed to result in a half-integer \( j \). The permanent number \( r = 2j_s + 1 \) of an even term becomes even. For an odd term, an integer \( j_s \) and an integer \( j_a \) are considered, as in the preceding paper, to give rise to an integer \( j \). The permanent number \( r = 2j_s + 1 \) of an odd term is odd. The multiplicity of a term turns out not to be larger than \( r \).

Sommerfeld puts Landé's g-factor in the form

\[
g = 1 + \frac{j(j + 1) + j_s(j_s + 1) - j_a(j_a + 1)}{2j(j + 1)}.
\] (7)

Landé's \( \tilde{J}^2 \) is equivalent to \( j(j + 1) \) in Sommerfeld's notation. Sommerfeld replaces Landé's \( R^2 \) and \( K^2 \) in Eq. (6) also by \( \tilde{R}^2 \) and \( \tilde{K}^2 \), getting the above formula of g-factor "symmetrical" for \( j, j_s \) and \( j_a \). Sommerfeld says that "the simplicity of the formula as well as the figures", with which the angles made by \( j_s \) and \( j_a \) are constructed in consideration of the values of \( j \) lying between \( |j_s - j_a| \) and \( j_s + j_a \) at intervals of unit magnitude, "seem to give evidence that our \( j, j_s \) and \( j_a \) are really the simplest elements for the description of facts".

When \( j \) takes its maximum value \( j_s + j_a \), the g-factor is given by \( g = \frac{(2j_s + j_a)}{j} \) and the splitting of the term becomes

\[
\Delta E = (2j_s + j_a)\hbar \Delta \nu_0 m/j,
\] (8)

where \( m \) is the magnetic quantum number, that is the component of \( j \) in the direction of the external magnetic field. Because \( \hbar \Delta \nu_0 = \mu_B H \), the part of the magnetic moment of the atom due to the unexcited state turns out to be proportional to the angular moment with the coefficient twice as large as the usual one. In the case of the s-term \( (j_a = 0) \), the magnetic moment of the atom becomes \( 2j_s \mu_B \), as has been pointed out by Landé. Sommerfeld says that this "reminds us of the magnetomechanical anomaly", and calls this factor 2 "the magneto-optical anomaly".

§3. The magnetomechanical anomaly

The "magnetomechanical anomaly" quoted by Sommerfeld is related to the problem of the origin of the magnetism of matter. It was well known that Ampère clarified the equivalence of a circular electric current with a magnetic
plate. The discovery of the electron induced the idea that the magnetism of matter would be caused by circular motions of electrons.

Richardson, in his 1908 paper "A Mechanical Effect Accompanying Magnetization",\(^{20}\) the relation between the magnetic moment and angular momentum of a magnetic substance arising from the motion of the constituent electrons in closed orbits. For a long thin bar of magnetic substance, he shows, taking into consideration the electrons rotating in planes perpendicular to the axis of the bar, that the relation

\[
\frac{U}{M} = \frac{2m}{e}
\]

holds between the magnetization \(M\) and angular momentum \(U\) per unit volume of the substance due to the motion of the rotating electrons (\(e/m\) being the specific electric charge of the electron). Richardson attempts to detect experimentally the occurrence of the angular momentum according to this relation in a long thin cylindrical bar of iron suspended by a fiber, by the sudden application of a vertical magnetic field. Richardson concludes, however, that "so far experimental success has not been attained owing to the difficulty of eliminating disturbing effects".

Einstein and de Haas consider, in their 1915 paper "Experimental Proof of Ampère’s Molecular Current",\(^{21}\) that "it results from the Curie-Langevin law that the magnetic moment of the molecule is independent of temperature, and exists therefore also at \(T = 0\), so that the energy of the motion of the circulating electrons would be a so-called 'zero-point energy'". From this viewpoint they measure the forced rotational oscillation by an alternating magnetic field of a cylindrical thin iron bar suspended in a coil, and arrive at the conclusion that the above relation "is also quantitatively confirmed at least approximately", though there is some experimental uncertainty.

According to a later report of de Haas,\(^{22}\) however, the experimental value of \(U/M\) is 14% smaller than the theoretical value.

Barnet discusses this problem in his 1915 paper "Magnetization by Rotation",\(^{23}\) saying that "it occurred" to him "while thinking about the origin of terrestrial magnetism". He makes a measurement of the magnetization of a body of magnetic substance which is given an angular acceleration. His

\(^{23}\)S. J. Barnet, Phys. Rev. 6 (1915), 239.
experimental result of \( U/M \) turns out to be smaller than the theoretical value by 56%. Stewart reports, in his 1918 paper "The Moment of Momentum Accompanying Magnetic Moment in Iron and Nickel",\(^{24}\) that the value of \( U/M \) is experimentally smaller than the theoretical value by 49% for iron and by 53% for nickel. Beck shows, in a 1919 paper "On the Experimental Proof of Ampère's Molecular Current",\(^{25}\) the result of an experiment similar to Einstein–de Haas's, giving values of \( U/M \) smaller than the theoretical one by 47% for iron and by 43% for nickel. Another similar experiment done in 1920 by Arvidsson\(^{26}\) brings about 1/2 of the theoretically expected value. These experiments show that the specific electric charge of the electron should be taken to be larger by a factor of about 2 than its known value, if the origin of ferromagnetism is attributed to the motion of rotating electrons.

Sucksmith and Bates report, in their 1923 paper "On a Null Method of Measuring the Gyro-Magnetic Ratio",\(^{27}\) the result of an experiment similar to Einstein-de Haas's but with an improved method. Sucksmith-Bates detect the magnetization of the iron rod by a small coil, from which the induced current is led to another coil through a resistance. The latter coil effects an inverse torque on a number of short magnetized needles, which are suspended from the rod by an aluminum wire. The resistance is adjusted so as to reduce the rotational oscillation of the rod driven by the alternating current to zero. They get \( U/M = (0.501 \pm 0.003) \times 2m/e \) for iron, nickel and Heusler alloys.

Sommerfeld notes that for the "magnetomechanical anomaly" "the factor 2 is determined within an error of 1% by Sucksmith and Bates". He attaches importance to its connection with the anomalous coefficient 2 of magnetic moment of the s-term according to Landé's indication,\(^{28}\) calling the anomaly of the coefficient of magnetic moment of the s-term "the magneto-optical anomaly".

Apart from ferromagnetic solids, for the magnetism of gases Weiss\(^{29}\) discovers in 1911 that the magnetization per mol is an integer multiple of 1123.5 c.g.s. This numerical value is much smaller than the magnetic moment per mol \( \text{Leh/}4\pi mc = 5584 \) c.g.s. (\( L \) being the Avogadro number), which results from the assumption that the electrons circulate in the Bohr orbits. Pauli

\(^{24}\)J. Q. Stewart, Phys. Rev. 11 (1918), 100.
\(^{28}\)Cf. §1.2 of the present section.
\(^{29}\)P. Weiss, Phys. Zeit. 12 (1911), 935.
treats this problem in his 1920 paper “Quantum Theory and Magnetron”.\(^{30}\)

Taking the directional quantization into consideration, he assumes that the angle \(\Theta\) between the external magnetic field and the angular momentum of an electron obeys the condition \(\cos \Theta = \pm m/k, m = 1, 2, \ldots, k\), and calculates the mean value of \(\cos^2 \Theta\), to which the magnetic susceptibility is proportional. He obtains good results matching with the experimental values for NO and O\(_2\) by putting \(k = 1\) and \(2\), respectively.

\section*{§4. Difficulties about the result of Stern-Gerlach’s experiment}

Stern and Gerlach in 1921 ~ 22 made a number of experiments,\(^{32}\) that are well known through text books on modern physics. In Stern-Gerlach’s experiment, a beam of atoms of vaporized silver from a furnace is observed, when it passes through an unhomogeneous magnetic field, to split into two components that fly onto a glass plate to form two separate stripes. Stern-Gerlach’s experiment is considered to show direct evidence for directional quantization. It shows also that the magnetic moment of the silver atom is of the order of the Bohr magneton \(\mu_B\). Sommerfeld describes in his book “Atombau und Spektrallinien” (4th ed., 1924) that “Stern and Gerlach have not only demonstrated \textit{ad oculos} the directional quantization of the atom in a magnetic field, but also proved the atomic nature of the magnetic moment”.

There are difficulties, however, in understanding the result of Stern-Gerlach’s experiment only in terms of the concept of directional quantization. As we have seen in Sec. 2, §1, Chap. 3 of the present volume, the directional quantization is discovered by Sommerfeld in 1916 through his application of the quantum condition to the azimuthal, zenithal and radial motions of the electron in the hydrogen atom, following a three-dimensional treatment due to Epstein and Schwartzschild of the Stark effect. According to Sommerfeld, as we have also seen there, the directional quantum number introduced by him “gives the first possibility of relating the polarization to the spatial position of the orbital plane”, and his way of calculation may be effective for Zeeman effect in which the magnetic field has axial symmetry. In Sec. 2, §7, Chap. 3 of the

\(^{30}\)W. Pauli, Phys. Zeit, \textbf{21} (1920), 615. In this paper, Pauli names \(\text{Leh}/4\pi mc\) the Bohr magneton. Nowadays, \(\mu_B = eh/4\pi mc\) is called the Bohr magneton.

\(^{31}\)Pauli “has excluded the value \(k = 0\) with Bohr”. (Cf. Sec. 2, §3 of the present chapter.)

present volume, it has been mentioned that the method of directional quanti-
ization is applied to the Zeeman effect by Debye and Sommerfeld independently
in 1916 with success.

In this way, the concept of directional quantization is understood in con-
nection with the variation of the frequency of a spectral line in the Zeeman
effect, that is, the splitting of spectral line. On the other hand, the light
from a luminous body is understood to be the radiation from atoms colliding
with each other, and the absorption spectrum to come from the transitions
of atoms to excited states under the effect of the radiation of the incident
light.

Thus, Einstein and Ehrenfest treat, in their 1922 paper "Quantum Theo-
retical Remarks on the Experiment of Stern and Gerlach", 33 the question "in
what way the atoms come to this orientation". Einstein-Ehrenfest point out
that "the atoms experience no collision while they are in the deflective mag-
netic field", and "they experience their last collisions in the vapor room of the
melting furnace". Now, "so long as one neglects the emission and absorption
of radiation and collisions or other similar influences", the atoms make the
motion of precession (Larmor rotation) around the direction of the magnetic
field. If the direction of the magnetic field changes slowly compared with the
speed of the motion of precession, the angle of precession remains unchanged.
Einstein-Ehrenfest stress that "an adjustment to the inclination demanded by
the quantum theory cannot therefore take place without external influences of
the kind of radiation or collisions".

On the other hand, "the most obvious explanation seems to be that the
adjustment of the atoms takes place by their entry in the field of the electro-
magnet, or, to be precise, through exchange of radiation". The time required
for the adjustment is estimated to be of the order of $10^{11}$ sec for a magnetic
field of 10,000 G, by taking "a corresponding classical model" of the magnetic
dipole that rotates conically and emits radiation, as "in the case of the tran-
sition from one to another quantum states". If one takes into account the
influence of the "positive and negative in-radiation" 34 of the surrounding ra-
diation at the room temperature, this time becomes of the order of $10^9$ sec.
These values of the time of the adjustment are too much longer than the transit
time $10^{-4}$ sec of the atoms through the electromagnet.

34) The "positive and negative in-radiation" correspond to the absorption and induced emis-
sion of radiation, respectively. [cf. Footnote 42)] in Sec. 3, §2, Chap. 3 of the present volume.
Einstein-Ehrenfest then suppose the following two cases for the interpretation of the result of Stern-Gerlach's experiment. In the first case, it is assumed that "in the vapor room of the melting furnace each silver atom is immediately quantized completely through collision". The magnetic axis of each atom lies in the direction of the magnetic field in the vapor room, even if the magnetic field is very weak. Part of the quantized magnetic moments is parallel, and the other part antiparallel to the magnetic field, and "their statistical distribution is controlled by the temperature and field strength in the vapor room". In this case, however, the magnetic axis of each atom should be considered to orientate in the direction of the magnetic field at every point in the vapor room, when each atom passes through the magnetic field of rapidly varying direction in the vapor room, as if in an adiabatic process in a magnetic field of slowly varying direction. "This demands a violation of the mechanical equations."

In the second case, it is assumed that in the vapor room "immediately after each collision the magnetic axis of an atom orientates arbitrarily with respect to the weak magnetic field existing there", not being quantized. The adjustment to the parallel or antiparallel orientation to the magnetic field "occurs through infrared radiation, or, to be precise, through out-radiation and positive and negative in-radiation", when each atom passes through the electromagnet. In this case, however, "the assumption is essential that such transitions from unquantized to quantized states correspond to transition probabilities of much larger order of magnitude than those in transitions from quantized to quantized states", "corresponding to an adjusting time of $10^{-4}$ sec". Einstein-Ehrenfest conclude, thus, "the difficulties listed above show how unsatisfactory are both the attempts of interpretation discussed here of the result found by Stern and Gerlach".

Against the estimation of the time required for the adjustment of each atom to the specified inclinations, made by Einstein-Ehrenfest with "a corresponding classical model", Bohr states a critical opinion in a footnote in Chap. 2, §4, "Correspondence Principle and Constitution of Radiation" of his review article "On the Application of the Quantum Theory to the Atomic Structure; I. The Fundamental Postulates of the Quantum Theory".\(^{35}\) Bohr says that "the influence of the field on the harmonic components" of the periodic motion in an atom, "not only consists in the addition of new oscillations, whose frequencies are proportional to the external force", that is, those of the Larmor type, "but

---

\(^{35}\)N. Bohr, Zeit, f. Phys. 13 (1923), 117.
also brings with itself a change of the harmonic components of motion already existing in the undisturbed atom". Bohr states that "the speed of adjustment of the atom in the field can therefore hardly be assessed from the lifetime of the stationary state of an imaginary atom, in whose motion only the former components exist, as it is attempted in the mentioned paper" of Einstein-Ehrenfest. He stressed that "for such speed of adjustment the lifetime of the stationary states of the excited undisturbed atom, on the contrary, is probably to be looked at as decisive".

On this opinion of Bohr, Smekal comments, in his paper "On the Directional Quantization in Magnetic Field",\(^{36}\) that "irregular orientations in magnetic field do not correspond to stationary states". Einstein-Ehrenfest's analysis is to be considered to indicate that the quantum condition bears a certain restriction as a in certain lawfulness a substantialistic stage. The quantum condition, when applied to a system making conditional periodic motions, gives the energies of the stationary states of the system, from which the frequencies of the spectral lines are calculated by means of Bohr's frequency relation. The quantum condition does not, however, furnish a dynamical law for the translational motion of an unexcited or excited atom in a magnetic field. Einstein-Ehrenfest's question concerning the interpretation of the result of Stern-Gerlach's experiment comes to be solved in the quantum mechanics, in terms of the superposition principle of wave functions.

Besides the difficulty analyzed by Einstein-Ehrenfest, the result of Stern-Gerlach's experiment includes another difficulty. Even if the result of this experiment is interpreted in terms of the directional quantization, the beam of vaporized silver atoms splits into two components, instead of three as is demanded by directional quantization. As Sommerfeld writes in his book cited above, "the behaviour of the silver atom in the magnetic field is not one of a hydrogen atom as in the normal Zeeman effect, but is of an anomalous kind as in alkalis". For each of the "magneto-mechanical anomaly", the "magneto-optical anomaly", and the anomaly of the atomic magnetic moment seen in Stern-Gerlach's experiment, it is refused to solve the problem of the factor "2".

2. The Shell Structure in the Configuration of Atomic Electrons

§1. The valence and the configuration of atomic electrons

Analyses of the anomalous Zeeman effect have clarified, as we have described in the last section, that the main parts of certain atoms exclusive of their respective luminous electrons, which are called atomic bodies (Atomrumpfe), atomic rests (Atomreste) or atomic cores, have anomalous coefficients of magnetic moment. The problem of the constitution of the atomic core itself from the atomic electrons, excepting the luminous electrons, was made clear through studies of the systematics of chemical properties of elements by means of the theory of atomic structure, which was developed with the spectroscopic properties of atoms as clues.

We have seen, in Sec. 4, §5, Chap. 2 of the present volume, that Bohr chose the number of the electrons in the outermost ring of the atom of an element so as to fit its valence, in the course of discussions of the configuration of atomic electrons on "the assumption of constancy of angular momentum", in the paper on his atomic model in 1913. We have also described, in Sec. 4, §6, Chap. 2, Vol. I, Thomson's attempt to explain the periodic law of elements on the basis of his atomic model. Though the number of atomic electrons was not determined in the latter atomic model, it became determined in the former atomic model. This opened the way to the clarification of correspondence between the valence and the configuration of atomic electrons.

Now, Kossel considers, in his 1916 paper "On Molecular Binding as Problem of Atomic Structure",\(^\text{37}\) that a neutral atom have the same number of electrons as the atomic number of the respective element, following the fact that van den Broek's assumption\(^\text{38}\) of the equality of the number of atomic electrons to the atomic number was confirmed by Mosley's experiment, which we have treated in Sec. 4, §6, Chap. 2 of the present volume. Kossel supposes that in atoms of rare gases "the firmness of electrons are so high that they are not stripped off in chemical substitutions". "The activity of valence in polar bond is understood as the interchange of a known number of electrons."

Concerning the force with which the charge of electricity due to interchange of electrons effects the binding of a polar molecule such as KCl, Kossel takes


The Shell Structure in the Configuration of Atomic Electrons

a simple model of two spheres of radius $10^{-8}$ cm, which are in contact and have at their centers point electric charges of opposite signs but equal magnitude $e$. He gets an estimation of $5.4 \times 10^{-12}$ erg for their mutual electrostatic energy, which is larger than the thermal energy $5.8 \times 10^{-14}$ erg at the room temperature. Kossel considers, however, that for molecules in water it should be divided by the dielectric constant of water to get $6.8 \times 10^{-14}$ erg, a value which may explain the dissolubility of KCl in water at the room temperature.

This Kossel’s idea is the one that gives a substantialistic ground to J. J. Bergelius’ theory at the beginning of the 19th century, in which molecular binding was assumed to come from the electric force between electrically positive and negative atoms. Kossel’s idea cannot explain, however, the molecular binding between atoms of the same kind such as those in $\text{H}_2$, $\text{N}_2$, etc. He says that “it may be expected that systems of such a kind could be understood” with the model\(^{39}\) treated in Bohr’s paper in 1913.

Kossel grasps substantialistically, on the basis of the Bohr model, the regularities of the valences in chemical bonds of molecules of various kinds which were analyzed by chemists R. Abegg, A. Wermer, et al. Kossel interprets the period 8 in the elements of atomic numbers from 1 to 25, by supposing that a neutral atom consists of an $n$-valent positive ion of a rare gas atom and $n$ electrons in a circular orbit around the ion ($n = 1, 2, \ldots, 8$). In this configuration of atomic electrons, $n$ electrons are supposed to be in the outermost circular orbit of the atom, $n$ being assumed to increase cyclically in order of the atomic number, differently from the configuration\(^{40}\) given by Bohr in 1913.

The importance of the number 8 of electrons in the configuration of atomic electrons is made clear, in the same year 1916, also by a chemist Lewis in his paper “The Atom and the Molecule”.\(^{41}\) Lewis assumes that the atomic electrons stay on concentric spherical shells. It is assumed that the innermost shell can contain at most 2 electrons, while every other shell 8 electrons. Since 8 electrons on a shell are supposed to take their positions at the corners of its inscribed cube, Lewis’ theory becomes known as “the octet theory”. He himself calls it “theory of the cubic atom”.

When there are $n$ electrons on the outermost shell of an atom, the respective element is considered to belong to the $n$th group in the periodic table

\[\text{Cf. Sec. 4, §5, Chap. 2, p. 99 of the present volume.}\]

\[\text{Cf. Sec. 4, §5, Chap. 2, p. 98 of the present volume.}\]

\[\text{G. N. Lewis, Journ. Amer. Chem. Soc. 49 (1916), 229.}\]
of elements. Lewis assumes that the outermost shell tends to hold an even number of electrons, especially 8 electrons. On this assumption, not only the polar binding such as in KCl is explained in the same way as in Kossel’s theory, but also the binding in O₂ molecule is explained similarly, by supposing that the cubes of the two O atoms are arranged to have a common side and the 12 electrons in total are at all the corners of the outermost cubes. In the case of F₂ molecule, the two cubes of the F atoms are assumed to have a common edge, so that the 17 electrons in total occupy all the corners of the two cubes. Langmuir names bindings of this kind “covalent bonds”.

According to the octet theory, Lewis can explain the chemical properties of 35 elements, among the known 85 elements, such as the rare gases, alkalis, alkaline-earth metals, halogens, and B, Al, Sc, C, Si, etc. Lewis’ octet theory should be said not to be concerned with the laws of the motion of atomic electrons, but to be concerned with the modes of the function of electrons in chemical binding of atoms, though it is based on the substantialistic knowledge that the number of atomic electrons is equal to the atomic number.

Of his own theory Lewis writes that “while it bears much resemblance to some current theories of the atom, it shows some radical points of departure from them”.

Langmuir makes, in his 1919 paper “The Arrangement of Electrons in Atoms and Molecules”, 42) “an extension of Lewis’ theory of the ‘cubical atom’”. Langmuir treats all the elements inclusive of rare earths, and discusses not only their chemical properties but also their physical properties such as boiling point, melting point, electric conductivity and so on. According to Rydberg’s indication 43) that the numbers of electrons in rare gas atoms are given in turn by \(2(1 + 2^2 + 2^2 + 3^2 + 3^2 + 4^2 + \cdots)\), Langmuir supposes that the atomic electrons are distributed through a series of concentric spherical shells. All the shells are assumed to be of equal thickness, and their effective radii stand in the ratio 1: 2: 3: 4, the effective radius of a shell being defined as the mean of its inner and outer radii. Each shell is divided into a number of cells of an identical volume, the first shell being divided into two cells. Each of the two cells of the innermost shell can contain only one electron, but each cell of the other shells is capable of holding two. It is thus assumed that the innermost shell, which is denoted by I, is of a single layer, but each of other shells consists of two layers, which are denoted in turn by II_a, II_b, III_a, III_b, IV_a.

43) J. R. Ryderg, Phil. Mag. 28 (1914), 144.
While the shell I is closed by H and He, the layer II	extsubscript{a} closes at Ne and the layer II	extsubscript{b} at Ar. The layer III	extsubscript{b} can contain 18 electrons, but because of the symmetry with respect to the 8 electrons in II	extsubscript{b}, the element Cu which corresponds to the configuration of 11 electrons in III	extsubscript{b} is assumed to return to the first group in the periodic group. The elements Fe, Co and Ni, which are represented by the configurations of 8, 9 and 10 electrons in III	extsubscript{b}, respectively, are known to have similar properties. It is supposed therefore that they can be taken together to endow III	extsubscript{a} with the period 8. The layers III	extsubscript{a} and IV	extsubscript{b} are considered in a similar way.

Langmuir writes that the conclusions arrived at in his theory "are not easily reconciled with Bohr's theory of the atom", and "Bohr's stationary states have a rather close resemblance to the cellular structure postulated" in the paper under consideration. Langmuir's theory is, Lewis' octet theory, like the one that is concerned with the modes of the function of electrons in chemical binding.

§2. Bohr's theory of atomic structure

Kossel's paper was published in the same year 1916 as Sommerfeld's paper on the spectra of hydrogen-unlike atoms, which has been treated in Sec. 1, §1 of the present chapter. Langmuir's paper was published in 1919, the same year as Kossel-Sommerfeld's paper, treated also in Sec. 1, §1 of the present chapter, in which they pointed out the "displacement law in series spectra". On the basis of the progress in the theory of spectra Bohr discussed, in his lecture given in Oct. 1921 "The Structure of the Atom and the Physical and Chemical Properties of the Elements", the configuration of atomic electrons, with the spectra and the physical and chemical properties of the elements as the clues.

One of the motives for this lecture was a letter to the editor of the journal "Nature" from a chemist Campbell. In this letter, Campbell stresses that, though Bohr-Sommerfeld's theory and Lewis-Langmuir's theory are said to be "mutually inconsistent in that the first indicates that the electrons are revolving in planetary orbits and the second that they are fixed in constant positions", "they are not really inconsistent". Campbell argues that Bohr-Sommerfeld's theory "assumes only that in stable states they have the energy,

\[^{44}\text{N. Bohr, Zeit. f. Phys. 9 (1922), 1 (German translation of the lecture); Essay III in "The Theory of Spectra and Atomic Constitution", Cambridge University Press, 1922 and 1924 (English translation of the lecture).}\]
\[^{45}\text{N. R. Campbell, Nature, Nov. 25 (1920), 408.}\]
calculated by Hamiltonian dynamics, which they would have if they were moving in certain orbits”, and “according to Bohr's principle of correspondence, the intensity and polarization of spectral components can be predicted by the application of classical dynamics to certain assumed orbits, although it must be assumed at the same time that the electrons are not moving in those orbits”.

In this way, Campbell avoids finding some consistency between the motion of electrons in Bohr-Sommerfeld's theory of spectra and the configuration of electrons in Lewis-Langmuir's theory of regularities of chemical bonds, and considers the concepts of the lawfulness of the motion of electrons in the former theory, such as the quantum condition, selection rule, transition probability, etc., simply as the methods of calculation. Consistency of this kind, however, should not be treated in such a plane-projected manner as arranging every thing on equal footing, but should be grasped in a stereo-structural manner which takes into consideration any appearance of the stratum structure of scientific cognition. This may be easily understood, if it is recalled that, in the history of the theory of the revolution of planets, the regularities in the Ptolemaic theory with the use of revolving circles had an accuracy higher than the laws in the Copernican theory.

Bohr tried, in a reply to Campbell, Mar. 24, 1921 issue of “Nature”, “briefly to indicate how it seems possible to overcome certain difficulties hitherto involved in the attempts to develop a general theory of atomic constitution based on the application of the quantum theory”. Bohr made a detailed discussion of this problem in the above-cited lecture in Oct. 1921.

In this lecture Bohr supposes that, for the neutral atom of an element of atomic number \( N \), “the process of formation may be regarded as occurring in \( N \) stages, corresponding with the successive binding of \( N \) electrons in the field of the nucleus”. This idea was already presented in his paper in 1913 on the atomic model. As Bohr says, however, Kossel\(^{46}\) “shows that a more detailed explanation of the origin of the high frequency spectra can be obtained on the basis of the group structure of the atom,” so that the existence of groups of electrons in the atom gets confirmed at the time of Bohr's lecture. In the lecture Bohr gives the opinion that each electron to be bound settles in one of the stationary states of the atom, which is specified by a pair of quantum numbers, the principal quantum number \( n \) and the azimuthal quantum number \( k \). Bohr calls the orbit of the electron in the stationary state specified by \( n \) and \( k \) the \( n_k \) orbit.

---

\(^{46}\)W. Kossel, Ref. 20) on p. 130, Sec. 2, §2, Chap. 3 of the present volume.
In the ground state of the hydrogen atom, the electron is in the $1_s$ orbit. As regards the formation of the helium atom, "we meet the much more complicated problem of the binding of a second electron". The spectrum of helium consists of two systems of lines, on account of which helium was once thought to be a mixture of two different gasses, called orthohelium and parhelium. But, according to Bohr, "now we know that the two spectra simply mean that the binding of the second electron can occur in two different ways".

The spectral terms of the parhelium series are singles, of which the lowest is a $S$-term, while those of the orthohelium series are triplets, of which the lowest is a $s$-term. Bohr says that, according to Landé and Kramers, the orthohelium spectrum is due to transitions between stationary states in which the two electrons move in the same plane, and the parhelium spectrum is ascribed to stationary states in which the planes of the two orbits form an angle with each other. On another side, Frank and his co-worker's experiments on inelastic scattering of electrons by helium atoms show that the first excited states are metastable. This, together with the selection rule $k \rightarrow k \pm 1$, shows that the ground state of the helium atom is a $1S$-term. From these considerations, Bohr concludes that the second electron of the helium atom, as well as the first, is bound in a $1s$ orbit. In metastable states, the second electron should therefore be in a $2s$ orbit. This corresponds to a firmness of binding which is about six times less than that for the electron in the ground state of the helium atom.

As regard the Li atom, its spectrum is hydrogen-like, and from experiments on selective absorption its ground state turns out to be a $S$-term, whose strength of binding is about a third of that of the hydrogen atom. From these considerations, Bohr concludes that the outermost electron is bound in a $2s$ orbit. He argues that if the third electron is in a $1s$ orbit together with the other two electrons, the process of binding of the third electron would be of a type entirely different from the transitions between the stationary states connected with the emission of the lithium spectrum.

About the elements Be, B and C, no such simple spectra are known that would guide us as in the case of Li. Bohr thus supposes, on general physical and chemical evidences, that the fourth, fifth and sixth electrons are in the $2s$ orbits, which are arranged in such mutual inclinations that their planes are parallel to the sides of a regular tetrahedron. Bohr expects that this would bring about a highly symmetrical configuration which might be capable of explaining the structure of organic compounds.
For elements from N to Ne in the second half of the second period, the seventh up to the tenth electrons are supposed to be bound in a circular $2_2$ orbit, on account of the symmetry reached by C. Bohr interprets the stability of Ne as a rare gas element, by introducing the assumption that “the planes of the four $2_2$ orbits must be regarded not only as occupying a position relative to one another characterized by a high degree of spatial symmetry, but also as possessing a configuration harmonizing with the four elliptical $2_1$ orbits”.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$n_k$</th>
<th>$1_1$</th>
<th>$2_12_2$</th>
<th>$3_13_23_3$</th>
<th>$4_14_24_34_4$</th>
<th>$5_15_25_35_45_5$</th>
<th>$6_16_26_36_46_56_6$</th>
<th>$7_17_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>He</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Li</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Be</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 B</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Ne</td>
<td>2</td>
<td>[44]</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Na</td>
<td>2</td>
<td>44</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Mg</td>
<td>2</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Al</td>
<td>2</td>
<td>44</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>A</td>
<td>2</td>
<td>44</td>
<td>[44]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>K</td>
<td>2</td>
<td>44</td>
<td>44</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Ca</td>
<td>2</td>
<td>44</td>
<td>44</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Sc</td>
<td>2</td>
<td>44</td>
<td>441</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Ti</td>
<td>2</td>
<td>44</td>
<td>442</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Cu</td>
<td>2</td>
<td>44</td>
<td>[666]</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Zn</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Ga</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Kr</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>[44]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Rb</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>44</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Sr</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>44</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Y</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>441</td>
<td>(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Zr</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>442</td>
<td>(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Ag</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>[666]</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Cd</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>666</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>In</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>666</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>X</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>666</td>
<td>[44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Cs</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>666</td>
<td>44</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>Ba</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>666</td>
<td>44</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>La</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>666</td>
<td>441</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>Ce</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>6661</td>
<td>441</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>Pr</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>6662</td>
<td>441</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Cp</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>[8888]</td>
<td>441</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>Hf</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>8888</td>
<td>442</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Au</td>
<td>2</td>
<td>44</td>
<td>666</td>
<td>8888</td>
<td>[666]</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Considering then the third period of the periodic table, Bohr supposes that in the Na atom the eleventh electron occupies a $3_1$ orbit in order not to destroy the symmetry reached by Ne. He makes this supposition for explaining the property of the Na atom that it is apt to become a positive ion. About the Mg atom, it is concluded from its known spectrum that the twelfth electron is in a $3_1$ orbit. Since the spectrum of Al shows that its first term is a $p$-term, the thirteenth electron should be concluded, on the other hand, to be in a $3_2$ orbit. Bohr mentions that this difference results from a special condition for the binding of the thirteenth electron. In the case of Si, Bohr assumes that the four outermost electrons occupy a $3_1$ orbit, with a symmetry similar to that of C in the second period.

Bohr proceeds to discuss elements in the fourth period, and supposes that in the K and Ca atoms the nineteenth and twentieth electrons, respectively, are in a $4_1$ orbit. He considers, however, that in the case of Sc an electron in a $3_3$ orbit is more strongly bound than in a $4_1$ orbit, because the electric charge of the nucleus is now so large that the difference between the forces inside and outside the orbit of the eighteenth electron becomes small. This is supported by the study of the spectrum of Sc. Bohr assumes in this way that "an inner group is being developed" in the fourth period. He supposes that this occurs also in the fifth, sixth and seventh periods.

The table in the foregoing page shows the configuration of atomic electrons given by Bohr. In this table, $N$ denotes the atomic number. The square brackets indicate the configurations at which the symmetry mentioned above is assumed to be reached, and the round brackets those for which the inner groups are assumed to develop.

\[\begin{array}{cccccccccc}
N \backslash n_k & 1_1 & 2_12_2 & 3_13_23_3 & 4_14_24_34_4 & 5_15_25_35_45_5 & 6_16_26_36_46_56_6 & 7_17_2 \\
80 \text{Hg} & 2 & 44 & 666 & 8888 & 666 & 2 \\
81 \text{Tl} & 2 & 44 & 666 & 8888 & 666 & 21 \\
86 \text{Em} & 2 & 44 & 666 & 8888 & 666 & [44] \\
87 & 2 & 44 & 666 & 8888 & 666 & 44 & 1 \\
88 \text{Ra} & 2 & 44 & 666 & 8888 & 666 & 44 & 2 \\
89 \text{Ac} & 2 & 44 & 666 & 8888 & 666 & 441 & (2) \\
90 \text{Th} & 2 & 44 & 666 & 8888 & 666 & 442 & (2) \\
\end{array}\]

\[\text{(47)}\] In this table, $N$ denotes the atomic number. The square brackets indicate the configurations at which the symmetry mentioned above is assumed to be reached, and the round brackets those for which the inner groups are assumed to develop.

\[\text{(48)}\] In the lecture under consideration Bohr presented a diagram, which he obtained by modifying a flow chart due to a chemist J. Thomsen at Copenhagen that expressed the
§3. X-ray spectra and structure of atoms

Bohr, in collaboration with Coster, attempts to get evidence for the development of the inner group mentioned above by analyzing X-ray spectra, in their paper "X-ray Spectra and the Periodic System of the Elements". As is described in Sec. 2, §2 of the foregoing chapter, in Sommerfeld’s analysis of X-ray spectral terms in 1916, he succeeded in explaining the X-ray doublet term by mean of the relativistic effect in the orbital motion of an atomic electron, but did not arrive at results in full accordance with the combination rule, since the selection rule was not yet known.

Owing to the discovery of the selection rule in 1918 and the development in experiments on X-ray absorption spectra, the analysis of X-ray spectral terms made great progress and it turned out that X-ray terms corresponding to electrons in the K, L, M, N, O and P shells with principal quantum numbers \( n = 1, 2, 3, \ldots \), are of multiplicities 1, 3, 5, 7, 5 and 3, respectively. With the use of the notation in Bohr-Coster’s paper, the triplet L-terms are denoted as \( L_1, L_{II} \) and \( L_{III} \), in the order of their depths, that is, in the order of the absolute values of the term energy. Hertz, after Moseley’s diagram, made a diagram of \( \sqrt{\nu/R} \) vs. the atomic number \( z \), \( \nu \) and \( R \) being respectively the wave number corresponding to the absolute value of the term energy and the Rydberg constant, and found that the difference \( \Delta \sqrt{\nu/R} \) in \( \sqrt{\nu/R} \) between the \( L_1 \) and \( L_{II} \) terms scarcely depends on \( Z \). On the other hand, about the \( L_{II} \) and \( L_{III} \) terms he found that the difference \( \Delta \lambda \) in their corresponding wave lengths is almost independent of \( Z \).

According to Sommerfeld’s relativistic formula of the energy level [Eq. (53) in Chap. 3], \( \nu/R \) is given by

\[
\frac{\nu}{R} = \frac{1}{n^2}(Z-s)^2 + \frac{\alpha^2}{n^4}(Z-s)^4 \left( \frac{n}{k} - \frac{3}{4} \right),
\]

periodic property of elements in terms of the atomic weight [Z. anorg. Chem. 9 (1911), 190]. Bohr used the atomic number, in stead of the atomic weight, and added marks to indicate the elements for which the development of the inner group was assumed. Since in Bohr’s later papers he showed several times this diagram without explicit reference to Thomsen, this diagram was often cited as Bohr’s diagram [cf. for example A. Rubinowicz, **Handbuch der Physik** (1933, Springer), Bd. 24]. This diagram should rather be called Thomsen-Bohr’s diagram.

50) G. Hertz, Zeit. f. Phys. 3 (1920), 19.
so that the relation \( \Delta \sqrt{\nu/R} = \text{const.} \) is considered to come from the difference of the screening constant \( s \) in the first term, while the relation \( \Delta \lambda = |\Delta \nu/\nu^2| = \text{const.} \) can be interpreted as due to the second term representing the relativistic effect, in the same way as Sommerfeld had suggested in 1916.\(^{51}\)

Wentzel\(^{52}\) confirmed, at Sommerfeld's institute, that a similar circumstance holds for the \( M \) and \( N \) terms, and called the pair of the fine terms of which \( \Delta \lambda = \text{const.} \) "regular doublet," and those for which \( \Delta \sqrt{\nu/R} = \text{const.} \) "irregular doublet". He introduced further a third quantum number \( m \), besides the principal and azimuthal quantum numbers \( n \) and \( k \), to distinguish these two kinds of doublet. He denoted the \( K; L_I, L_{II}, L_{III}; M_I, \ldots \) terms in terms of the pair of the quantum numbers \( (k, m) \), as \((1, 1); (1, 1), (1, 2), (2, 2); (1, 1)\ldots \) respectively, of which \( m \) was assumed to take the values \( m = k \) and \( k + 1 \). Wentzel named \( m \) "the grand quantum number," and supposed the selection rules \( m \rightarrow m \pm 1 \) and \( k \rightarrow k, k \pm 1 \) to hold. A quantitative analysis of X-ray terms was further made with this method by Sommerfeld and Wentzel.\(^{53}\)

Coster\(^{54}\) made at Siegbahn's institute an analysis of similar X-ray spectra and introduced two kinds, \( a \) and \( b \), of X-ray terms, and assumed that transition could not occur between two terms of an identical kind. Coster\(^{55}\) says that "we do not yet know the physical meaning of this grand quantum number," and names Wentzel's irregular doublet "screening doublet," arguing that "as the screening doublets show no irregularity at all" Wentzel's naming "seems not to be well chosen". Coster names Wentzel's regular doublet "relativity doublet," for the reason that this doublet is due to the relativistic effect represented by the second term of the formula shown above. Coster concludes, from the diagram of energy levels obtained from the X-ray terms, that his results "are in beautiful agreement with Bohr's theory as regards successive development of the shells of electrons in the atom".

Since Bohr's theory of atomic structure presumes the \( n_k \)-orbit of the atomic electron, it becomes necessary to introduce some way of distinguishing the two kinds of X-ray terms. Thus, in Bohr-Coster's paper cited above, the X-ray terms are designated with the set of quantum numbers \( n(k_1, k_2) \). Here, \( n \) is the principal quantum number, and \( k_1 \) and \( k_2 \) are the same as Wentzel's \( m \)

\(^{51}\)Cf. Sec. 2, §4, p. 130, Chap. 3 of the present volume.

\(^{52}\)G. Wentzel, Zeit. f. Phys. 6 (1921), 84.

\(^{53}\)A. Sommerfeld und G. Wentzel, Zeit. f. Phys. 7 (1922), 86.

\(^{54}\)D. Coster, Zeit. f. Phys. 6 (1921), 185.

\(^{55}\)D. Coster, Phil. Mag. 43 (1922), 1070; 44 (1922), 546.
and $k$, respectively, $k_2$ being supposed to take the values $k_2 = k_1$ and $k_1 - 1$. However, Bohr-Coster "conceive their theoretical significance in a somewhat different manner". Bohr-Coster assume that $k_1$ and $k_2$ obey the selection rules $k_1 \to k_1 \pm 1$ and $k_2 \to k_2, k_2 \pm 1$.

In Bohr-Coster’s paper, the $K; L_1, L_{II}, L_{III}; M_1, M_{II}, M_{III}, M_{V}; \ldots$ terms are denoted as $1(1, 1); 2(1, 1), 2(2, 1), 2(2, 2); 3(1, 1); 3(2, 1), 3(2, 2), 3(3, 2), 3(3, 3); \ldots$ respectively. This amounts to the assumption that $k_1$, on one hand, is the azimuthal quantum number which specifies the difference in the screening constant $s$ due to the difference of the orbit and is independent of the second term in the above formula, while $k_2$, on the other hand, is the azimuthal quantum number which gives the difference in the relativistic effect due to the difference of the orbits, and is independent of the screening effect.

Bohr-Coster say that "no complete explanation can be given, at the present date of the theory, for the appearance of these doublets". They call energy levels for which $k_1 = k_2$ "normal levels," and those for which $k_1 \neq k_2$ "abnormal levels". By supposing that "each of these (abnormal) levels is quantitatively closely related to two normal levels" and "appears in the atom together with a normal level," Bohr-Coster consider that Bohr’s theory of the structure of atoms based on $n_k$-orbits is confirmed also from the analysis of X-ray spectra. They emphasize in particular the fact that the lines of $\sqrt{\nu/R}$ vs. $Z$ of the $N$ terms and of the $O$ terms in a Moseley-like diagram deviate from the straight and break or cross each others near those $Z$’s for which the inner groups have been assumed to develop.

In Bohr-Coster’s paper, however, no mention is made of what kind of orbit the abnormal level correspond to. Even if the normal level is considered to correspond to $n_k$-orbit with the azimuthal quantum number $k \equiv k_1 = k_2$, it remains vague whether the abnormal level corresponds to $n_{k_1}$-orbit or $n_{k_2}$-orbit, or something else. This point is treated later, as we shall see below, by Landé and by Millikan and Bowen.\textsuperscript{56}

\textsuperscript{56} J. Mehra and H. Rechenberg include this subject in “Atomic Structure and the Problem of Closed Electron Shells,” VI-2, Part 2, Vol. I in their book “The Historical Development of Quantum Theory” (1982, Springer-Verlag). Although they treat Wentzel’s and Coster’s works in parallel ways, they do not make, in this connection, any mention of Bohr-Coster’s paper. If one does not see the difference between the considerations of Sommerfeld–Wentzel and of Bohr-Coster noted above, it would be impossible to understand the significance of Landé’s and Millikan-Bowen’s works.

It is utterly incomprehensible that Mehra and Rechenberg do not treat Bohr-Coster’s paper in connection with the subject under consideration. As will be seen shortly later, the “ground quantum number” introduced by Wentzel is discussed in Sommerfeld’s famous
About "the ground quantum number" Sommerfeld says, in his book "Atombau und Spektrallinien," 3'rd Edition (1922), that "we define a third quantum number, which we would like best to call 'inner quantum number' following the optical spectra, but on which we prefer, in view of the actual incomplete analogy with them, to impose the less connected name 'ground quantum number'". Landé points out in a 1923 paper the possibility that this similarity may hold good. He takes notice of the fact that the selection rules $\Delta k_1 = \pm 1$ and $\Delta k_2 = 0, \pm 1$ are the same as the selection rules $\Delta k = \pm 1$ and $\Delta j = 0, \pm 1$ for the azimuthal quantum number $k$ and inner quantum number $j = k, k - 1$ in the optical spectrum. He thus considers that the "relativity doublets" such as $(L_{II}L_{III}), (M_{II}M_{III}), (M_{V}M_{V})$, and so on, correspond to doublets in the optical spectrum. Landé gives attention to this similarity between the selection rules, by supposing that "the 'anomalous' $n(k_1, k_2)$-term arises from the removal of either a $n_{k_1}$-electron or a $n_{k_2}$-electron from the atom".

Landé shows that the difference in wave number $\Delta \nu = np_1 - np_1$ of the optical doublet $(np_1, np_2)$ is nearly proportional to $(Z - 9)^4$ for $n = 3, 4, 5$ and 6 in the case of Na, Mg$^+$ and Al$^{++}$. From the values of this difference he infers inversely those of the difference in wave number of the "relativity doublet" $(L_{II}L_{III})$. He gets a result which is in order-of-magnitude agreement with experimental findings, by changing the screening constant from 9 to 7 in consideration of the number of electrons occupying the orbits up to the $2s$-orbit. Similar results are obtained also for the relativity doublets $(M_{II}M_{III})$ and $(N_{II}N_{III})$, the screening constant being taken to be 19 and 21 respectively.

On making the relation $k_2 = k_1, k_1 - 1$ correspond to the relation $j = k, k - 1$, Landé adds that "according to recent ideas $j$ should be $k \pm 1/2$" in a footnote, to call our attention to the half-integer quantum number. Bohr calls into question the half-integer quantum number, on the other hand, in a footnote to a section on the anomalous Zeeman effect in a paper in the same year 1923,
"Line Spectra and Atomic Structure,"\textsuperscript{58} saying that "the introduction of half-integer \( j \)-values in the formal representation of complex structures tried in the text gives expression merely to the view that in these problems at the present states of the theory no direct connection can be achieved with the verified features of the quantum theory of periodic systems". Since in Bohr's theory importance is attached to the correspondence between the Fourier development of the periodic motion and the quantum nature of atomic systems, it would be difficult to include half-integer quantum numbers into the framework of his theory.

As we have described in Sec. 1 of the present chapter, Landé expressed by half-integer quantum numbers the fact that, in the anomalous Zeeman effect, a doublet term of an alkali atom splits into \( 2(2k - 1) \) states in total. Of this number of states Bohr makes the following discussion in his paper under consideration. He first argues that the core of an alkali atom can take on one position in an external magnetic field. This is because he explains the lack of paramagnetism in the noble gases, whose atoms are identical with the cores of the alkali atoms, by assuming that "the atoms of these elements can appear with their angular momenta only at right angles to the direction of the field". In Bohr's theory of the atomic structure, the closed \( K \)-shell has a finite amount of angular momentum differently from the other closed shells. Bohr nextly argues that the valence electron in the \( n_k \)-orbit can take on \( 2k \) positions in a magnetic field, because he "excludes the position, in which the angular momentum is perpendicular to the field, and which in the limiting case, when the central orbit turns into a Kepler orbit, should be regarded as singular".

The number of states of an alkali atom in a magnetic field becomes \( 2k \), that is only about half the observed number \( 2(2k - 1) \) even for large \( k \). Bohr thus assumes that "because of stability properties of the atom which cannot be described mechanically, the coupling of the valence electron to the atomic core is subject to a compulsion (\textit{Zwang}),... which means that to the atomic core, instead of the single possibility of adjustment in a constant external field, two different positions in the atom are forced, while on the other hand according to the same compulsion the outer electron, instead of the \( 2k \) possibilities of adjustment in an external field, holds only \( 2k - 1 \) possibilities of adjustment in the atomic connection".

\textsuperscript{58}N. Bohr, Ann. d. Phys. \textbf{71} (1923), 228.
§4. Stoner’s revision of Bohr’s configuration of atomic electrons

Of the optical doublet $2p$-terms, namely the $(2p_1, 2p_2)$’s of Li, Be$^+$, B$^{2+}$ and C$^{3+}$ atoms, which have the identical number of electrons but are different in the atomic number, Landé calculates in his paper in 1924 “The Essence of the Relativistic Röntgen Doublet”\(^{59}\) the screening constant from the difference in wave number, by supposing that these optical doublets terms should be relativity doublet terms, and gets a value nearly equal to 2. For Na, Mg$^+$ Al$^{2+}$ and Si$^{3+}$ atoms, he makes a similar calculation, getting a value of about 9 for their screening constant. This shows that there holds in fact a similarity between the optical and relativity doublet terms.

Independently of Landé, Bowen and Millikan analyze in their paper “The Extension of the X-ray Doublet Laws into the Field of Optics”\(^{60}\) many optical doublet terms, inclusive of those in the ultraviolet region. They calculate the screening constants from the differences in wave number of the doublet $2p$-and $3d$-terms of the set Li$^+$, Be$^{2+}$, B$^{3+}$, C$^{4+}$ and N$^{5+}$, of the doublet $3p \sim 6p$, $3d$-, $4d$- and $4f$- terms of the set Na$^+$, Mg$^{2+}$, Al$^{3+}$, Si$^{4+}$, and P$^{5+}$, and of others, showing that they are the regular doublet terms (relativity doublets in Coster’s naming). They deduce also the differences in the square root of wave number of the pairs of terms $(2s, 2p_2)$ of the set Li$^+ \sim N^{5+}$, and of the pairs of terms $(3s, 2p_2)$ and $(3p_1, 2d_2)$ of the set Na$^+ \sim P^{5+}$, concluding that they are the irregular doublet terms (screening doublet terms in Coster’s naming). They state that “the behavior, then, of the irregular doublet is beautiful confirmation of the conclusion arrived at from the behavior of the relativity doublet that both of the doublet laws first discovered in the X-ray field are equally applicable to the whole optical region”.

On the basis of this result, Millikan and Bowen point out in their next paper “Some Conspicuous Successes of the Bohr Atom and a Serious Difficulty,”\(^{61}\) that the configuration of atomic electrons in Bohr’s theory of atomic structure should be revised. The X-ray terms $L_1$, $L_{II}$ and $L_{III}$ are now identified with the optical terms $s$, $p_2$ and $p_1$, respectively. The terms $s$, $p_2$ and $p_1$ are therefore concluded to correspond to the $2_1$-, $2_1$- and $2_2$-orbits, respectively. It thus results that the $p_2$- and $p_1$-orbits have one and the same screening constant,

\(^{60}\)I. S. Bowen and R. A. Millikan, Phys. Rev. 24 (1924), 209.
in spite of the difference between their shapes, while the $s$- and $p_1$-orbits have different screening constants from each other in spite of the sameness of their shapes. Millikan-Bowen say that "to avoid these two difficulties it is necessary to make the very unnatural assumption that the inner quantum number, i.e., the orientation of the orbit with respect to the atomic body, has a much more powerful effect upon the screening constant $s$ than has the wide difference in the shape of the orbit."

Millikan-Bowen state that these results "force a choice between one or the other of the two horns of a dilemma: (1) The abandonment of relativity causes and effects altogether as regards electronic orbits, or (2) the abandonment of Bohr's interpretation ideas and with them the practice of assigning azimuthal quantum numbers 1, 2, 3, 4 etc. to $s$, $p$, $d$, $f$, etc. terms, respectively".

Since in Millikan-Bowen's work the optical doublet term is interpreted on the basis of Sommerfeld's relativistic formula, the value of $k$ of the $p_2$- and $p_1$-terms should be different from each other. They therefore make the $p_2$- and $p_1$-terms correspond to the $2_1$- and $2_2$-orbits, respectively, confronting with the dilemma mentioned above. In Bohr-Coster's paper, on the other hand, such a dilemma does not arise because two kinds of azimuthal quantum numbers are used in turn. Since, however, the X-ray and optical doublet terms are now found to be identical with each other and the latter term is determined by the quantum numbers $k$ and $j$, it is no more of use to go on with Bohr-Coster's vague idea of two kinds of azimuthal quantum numbers $k_1$ and $k_2$. In fact, Millikan-Bowen say in a footnote that "the only escape from one or the other of these conclusions is to discredit the foregoing evidence for the correspondence between the $s$ and $L_1$ levels and between the $p_2p_1$ and the $L_{II}L_{III}$ levels".

Landé classifies the $L_1$, $L_{II}$ and $L_{III}$ terms, or the $s$-, $p_2$- and $p_1$-terms, using pairs of the quantum numbers $(k, j)$ as $(1, 1)$, $(2, 1)$ and $(2, 2)$ respectively. In order to explain the difference in energy of the regular doublet terms $L_{II}$ and $L_{III}$, or $p_2$ and $p_1$, by the difference in $j$ instead of $k$, Landé calculates in his paper cited just above,\textsuperscript{62} the average value of the magnetic field which the electron produces at the center of its orbit, as Heisenberg did before.\textsuperscript{63} It is assumed that the atomic body (atomic core) at the center of the orbit has an angular momentum, which can take parallel or anti-parallel orientation to that of the electron. For the difference in the magnetic interaction energy

\textsuperscript{62} A. Landé, Ref. 59) in the present subsection.

\textsuperscript{63} Cf. Sec. 1, §2 of the present chapter.
arising from these two possibilities, Land gets a result which is proportional to 
\((Z - s)^3\), but not to \((Z - s)^4\).

By the way, “the abandonment of relativity causes and effects” in Millikan-Bowen’s saying is, as we shall see in the next section, in relation to the process in which instead of Sommerfeld’s relativistic formula the spin-orbit interaction due to relativistic kinematics is introduced, and then Dirac’s relativistic wave equation inherently including the spin of the electron is established.

Now, taking notice of Landé’s papers, Stoner says in his paper “The Distribution of Electrons among Atomic Levels,” that “Lande shows that the relativity and optical doublet separations can both be represented by the same general formula, and places beyond doubt that the two types of doublet are essentially similar in origin,” and, making use of the azimuthal as well as inner quantum numbers \(k\) and \(j\), revises Bohr’s configuration of atomic electrons. Following Landé, Stoner classifies in terms of \((k, j)\) the sublevels \(K, L_1, L_{11}, L_{111}, M_1, \ldots\), of the atomic levels \(K, L, M, \ldots\).

“Without laying too much stress on any definite physical interpretation, or pressing analogy too far,” Stoner suggests that “for an inner sublevel the number of possible orbits is equal to twice the inner quantum number, these orbits differing in their orientation relative to the atom as a whole”. He assumes that “electrons can enter a group until all the possible orbits are occupied, when the atom will possess a symmetrical property”. In this way, the number of elections in a sublevel, which in Bohr’s theory of atomic structure has been assigned in consideration of the valence and other chemical properties, turns out to be determined by the inner quantum number. If the inner quantum number \(j\) is given the value \(j = k\) or \(j = k - 1\), twice the inner quantum number gives the observed term multiplicity as revealed by the spectra in a weak magnetic field, as has been clarified through the investigations of Landé and others, described in Sec. 1 of the present chapter, with the use of the half-integer quantum number. The following table shows the number, due to Stoner, of electrons in atomic levels.

Although Stoner considers that “there are few properties which depend solely on the number of electrons in levels and not also on the type of levels and its energy value,” he investigates intensities of X-ray lines, absorption of X-rays, chemical and magnetic properties, and optical spectra, concluding that “evidence is shown to give considerable support to a distribution of the

\[64\)E. C. Stoner, Phil. Mag. 48 (1924), 719.\]
kind put forward". For example, he explains the fact that the ratio of the intensity of the $\alpha_1$ X-ray to that of the $\alpha_2$ X-ray is equal to 2/1 from Fe to W by the supposition that there are twice as many electrons in the $L_{II}$ as in the $L_{II}$ sublevel.

### 3. The Spin of Electron and the Exclusion Principle

#### §1. Pauli’s analysis of the anomalous magnetic moment of the K-shell and the double-ness of the electron

As has been described in Sec. 1, §2 of the present chapter, Landé and Sommerfeld showed that the atomic magnetic moment in the $s$-terms is given by $2j_s\mu_B$, in terms of the inner quantum number in unexcited state $j_s$ and the Bohr magneton, so that the coefficient of magnetic moment takes the value 2 instead of the value 1 given by the classical theory. In Bohr's theory of atomic structure, on the other hand, the closed K-shell has a nonzero angular momentum, differently from the other closed shells. This implies that every rare gas atom in its unexcited state has the anomalous coefficient of magnetic moment 2.

Pauli points out, however, in his paper "On the Influence of the Velocity Dependence of the Electron Mass on the Zeeman Effect," that for atoms with large atomic number $Z$ the Bohr magneton $\mu_B$ should be multiplied by a correction factor $\gamma$, which is equal to the time average of $\sqrt{1 - \frac{u^2}{c^2}}$, because the relativistic correction becomes necessary for the electron mass. For an electron in the K-shell $\gamma$ is given by $\gamma = \sqrt{1 - \alpha^2z^2}$ in terms of the fine structure constant $\alpha$. The value of $\gamma$ is smaller than 1 by about 7% for $Z = 56$ and about 18% for $Z = 80$. Pauli studies the influence of this relativistic correction on the anomalous Zeeman effect.

---

65) W. Pauli, Zeit. f. Phys. 31 (1925), 373.
Pauli supposes that the atom consists of the atomic body and an outer electron, and denotes the coefficient of magnetic moment of the atomic core generally by $\kappa$. The change in energy of the atom in a weak external magnetic field is given by Eq. (3) on p. 182 together with the Landé's $g$-factor

$$g = 1 + (\kappa - 1)\delta,$$

(10)

where $\delta$ is given by

$$\delta = \frac{j(j + 1) + j_s(j_s + 1) - j_a(j_a + 1)}{2j(j + 1)},$$

(11)

according to Sommerfeld. Equations (10) and (11) with $\kappa = 2$ coincide with Eq. (7) on p. 189.

If the angular momentum of the closed $K$-shell is not zero, the coefficient of magnetic moment $\kappa$ of the atomic body of an alkali atom should be put, in consideration of the relativistic correction, to be $\kappa = 2\gamma$, instead of $\kappa = 2$. Or, in another way, if the anomalous coefficient of angular momentum $\kappa = 2$ is supposed to result formally from the addition of 1 to the classical value 1, $\kappa$ should be put to be $\kappa = 1 + \gamma$. For the outer electron $\gamma$ can be taken to be $\gamma = 1$. Pauli gets thus for the alkali atom

$$\kappa = 2\gamma, \quad g = 1 + (2\gamma - 1)\delta,$$

(12a)

$$\kappa = 1 + \gamma, \quad g = 1 + \gamma\delta.$$  

(12b)

For an alkaline earth atom, if one makes the additional assumption that the value of $\kappa$ of its atomic body coincides with the value of $g$ of the $s$-term of the alkaline-earth ion, one gets the same formula as given just above.

The $g$-factor in Eq. (12a) or Eq. (12b) depends systematically on $Z$, and is not compatible with observations. Pauli thus concludes that “if one holds on to the assumption that closed groups of electrons and in particular the $K$-shell are the origin of the magneto-mechanical anomaly, one must accept, not only a doubling of the ratio of magnetic moment to angular momentum compared with its classical value, but also a compensation of the relativistic correction”. He says that “the assumption of such a compensation is indeed in the present state of the theory logically possible, since the accepted doubling of the ratio under consideration is an ad hoc hypothesis unfounded closely”. Pauli is, however, “rather inclined to doubt the correctness of the assumption
of the essential participation of the noble gas configuration in the atom to the complex structure and the anomalous Zeeman effect, in the form of the angular momentum of the atomic body with anomalous magnetism”.

He then makes the following arguments. Though, if the angular momentum of the closed $K$-shell is assumed to be different from zero, while those of the other closed shells are assumed to be zero, the $K$-shell would be outstanding over the latter shells in some way, no counterpart of this configuration of electrons exists in the observable properties. For these properties, including the diamagnetic behaviour of the noble gases, it seems most natural to assign a vanishing angular momentum for all the closed shells, as was proposed by Sommerfeld. If one assigns to the atomic core a value of the ratio between the magnetic moment and angular momentum other than that for the lighting electron, one meets with the difficulty that as soon as the lighting electron is bound at last equivalently with the electrons in the atomic core, the value of this ratio should change suddenly. Furthermore, in consideration of the results due to Landé and Millikan-Bowen, it seems that the energy difference of an alkali doublet term can hardly be attributed to the deviation of the force of the atomic core on the lighting electron from the spherical symmetry, and to various orientations of the orbital planes of the lighting electron relative to the atomic core.

Through these analyses, Pauli arrives at the following conclusion — “The closed configurations should not contribute to the magnetic moment and to the angular momentum of the atom. Especially in the case of the alkalis, the angular momentum of the atom and its energy changes in an external magnetic field are considered essentially as a self effect of the lighting electron, which will be also regarded as the seat of the magneto-mechanical anomaly. The doublet structure of the alkali spectra as well as the breakdown of the Larmor theorem, takes place through a characteristic and classically indescribable kind of double-ness (Zweideutigkeit) of the quantum-theoretical property of the lighting electron”.

In this way, Pauli treats the atomic core, which was introduced substantially as the being to share the orbital angular momentum of the lighting electron and was supposed to have the anomalous coefficient of magnetic moment in its relation to the closed shells in the theory of atomic structure, making clear that the very attribute ascribed to the atomic core should belong

---

66) A. Landé, same as Ref. 59) in the foregoing section.
67) R. A. Millikan and I. S. Bowen, same as Refs. 60) and 61) in the foregoing section.
to the lighting electron itself, through the analysis of the contradiction in the notion of the atomic core.

§2. The shell structure of atoms and the exclusion principle of electrons

Pauli analyzes, in his paper “On the Connection of the Closure of the Electron Group in the Atom with the Complex Structure of Spectra,” received six weeks after the paper cited above, the atomic shell structure and the multiplicity of spectral lines on the basis of the “double-ness” of the property of electron mentioned above. He uses alkali atoms as the starting point of his analysis and bases his study on “the Zeeman effect of the alkalis as the fact of experience”.

Pauli denotes the azimuthal quantum number by \( k_1 \), and discriminates the alkali doublet term by \( k_2 = k_1 \) and \( k_2 = k_1 - 1 \), with only \( k_2 = k_1 \) for \( k_1 = 1 \). Pauli’s \( k_2 \) corresponds to the inner quantum number. He lets \( j \) be the total angular momentum of the atom, as in Sommerfeld’s paper in 1924. Its component in the direction of the external magnetic field is denoted by \( m_1 \). \( j \) is given by \( j = k_2 - 1/2 \) in terms of \( k_2 \), and the number of states in the magnetic field by \( 2j + 1 = 2k_2 \). The total number of states of the doublet in the magnetic field is equal to \( 2(2k_2 - 1) \).

In the case of the alkaline earths, there exist singlet and triplet terms, corresponding to the numbers of states in the magnetic field \( 1(2k_1 - 1) \) and \( 3(2k_1 - 1) \) respectively. This has been considered to arise for the reason that the lighting electron has \( 2k_1 - 1 \) ways and the atomic core has 1 and 3 ways, of orientation of the angular momentum. These numbers, 1 and 3, are different, however, from the number 2 assumed for the ways of orientation of the free atomic core of an alkaline earth atom, to which the s-term of an alkali atom corresponds. Pauli states, in opposition to the idea of non-mechanical “compulsion” proposed in Bohr’s paper in 1923, that “we can now simply interpret the total \( 4(2k_1 - 1) \) states of the atom as meaning that to the atomic core two states still belong in the field and to the lighting electron \( 2(2k_1 - 1) \) states as in the case of the alkalis”.

On the assumption that not only the numbers of states of the atomic core and of the lighting electron in a strong magnetic field, but also their energies are

---

68) W. Pauli, Zeit. f. Phys. 31 (1925), 765.
69) A. Sommerfeld, Ref. 18) in Sec. 1, §2 of the present chapter.
70) N. Bohr, Ref. 58) in Sec. 2, §3 of the present chapter.
additive, the components in the direction of the field of the angular momentum and the magnetic moment of an alkaline earth atom, \( \overline{m}_1 \) and \( \overline{m}_2 \) respectively, are given by summing up the respective quantities of the two valence electrons of the atom. Taking the values \( m_1 = \pm 1/2 \) and \( m_2 = \pm 1 \) for the components of the angular momentum and the magnetic moment, respectively, of each valence electron, which are taken to be the same as the values assumed for the \( s \)-term of the alkalis, one gets for the atom as a whole \( \overline{m}_1 = -1, 0, 0, 1 \) and \( \overline{m}_2 = -2, 0, 0, 2 \). This explains the \( s \)-terms of the alkaline earths, that is, the singlet \( S \)-term and the triplet \( s \)-term. These terms correspond to the terms with \( j = 0 \) and \( j = 1 \) of the atom in a weak magnetic field. To obtain the \( p \)-, \( d \)-, \ldots terms of the atom, it suffices to suppose that one valence electron is in the state of the \( S \)-term and the other valence electron takes the same values of \( m_1 \) and \( m_2 \) as those in the alkali \( p \)-, \( d \)-, \ldots terms, respectively.

In order to relate the number of stationary states of an atom in a strong magnetic field to that in a weak magnetic field, as is done above, Pauli makes use of the adiabatic theorem. According to this theorem, the number of stationary states of an atom in a magnetic field remains invariant when the strength of the magnetic field is adiabatically varied, so that from the addition rules \( \overline{m}_1 = \Sigma m_1 \) and \( \overline{m}_2 = \Sigma m_2 \) in the case of a strong magnetic field, the number of the terms and their quantum numbers in the case of a weak as well as vanishing magnetic field can be known. As Pauli stresses, however, the term energies or their intervals cannot be predicted by means of this theorem alone.

Now, according to Stoner,\(^{71}\) the number of electrons in a closed subgroup with the azimuthal quantum number \( k_1 \) is \( 2(2k_1 - 1) \), and this number coincides with that of the stationary states with the quantum number \( k_1 \) of an alkali atom in a magnetic field. Pauli considers that "we can make precise and generalize Stoner's idea" by considering the case of a strong magnetic field, and attributes Stoner's result to the following "general rule of the occurrence of equivalent electrons in the atom". "Equivalent electrons" here mean "electrons of the same value completely in regard to their quantum number as well as their binding energies".

The "general rule" says: "There can never be two or more equivalent electrons, for which the values of all the quantum numbers \( n, k_1, k_2, m_1 \) (or, what is the same, \( n, k_1, m_1, m_2 \)) coincide in a strong field. If an electron exists in an atom for which these quantum numbers have (in an external field) definite

\(^{71}\) C. Stoner, the same as Ref. 59) in Sec. 2, \$4 of the present chapter.
values, this state is full”. Pauli says that “we cannot give any closer reason for this rule, but it seems to appear of itself very naturally”.

He takes some simple cases to see whether conclusions from the rule agree with the experience or not. Firstly, from the rule the periods 2, 8, 18, 32, … the periodic system of elements are immediately obtained. This is because the number of equivalent electrons of given values of \( k_1 \) and \( k_2 \) cannot be larger than those of the allowed values of \( m_1 \) (namely \( 2k_1 \)), and just one electron belongs to each value of \( m_1 \) in the closed shells. Secondly, it follows that in the case of the alkaline earths no such triplet \( s \)-term exists that has the same principal quantum number as that of the ground state. In this case, for the equivalent binding of two electrons in a \( s \)-term to occur, two electrons should take the values \( m_1 = \pm 1/2 \) oppositely to each other in a strong magnetic field, so that \( m_1 \) can take only the value 0. Therefore, in a weak or vanishing magnetic field also only the value \( j = 0 \) is possible (singlet \( S \)-term).

Pauli considers further the case in which one electron is removed from a closed shell, as occurs in X-ray spectra. He names the value of \( m_1 \) of that electron “lacking value” of \( m_1 \). The value of \( m_1 \) of the whole atom is oppositely equal the lacking value of \( m_1 \). The spectrum of a singly ionized closed shell therefore corresponds to that of an alkali, as experience shows. He says that this is a special case of a general “reciprocity law”. To every configuration of electrons, he considers its “conjugate configuration” in which the lacking values of \( m_1 \) and the filled values are interchanged. The values of \( j \) of mutually conjugate configurations are identical. Because of this reciprocity law, the situation at the end of a period of the natural elements is attributed to that at the beginning of a period.

Pauli then discusses, the formation of an eight-electron shell (in which no electron with \( k_1 > 1 \) of a given principal quantum number \( n \) exists in the ground state). The subgroup of electrons with \( k_1 = 1 \) is assumed to be closed. The case of two electrons with \( k_1 = 2 \) is the same as that of the alkaline earths treated above. By adding the remaining six electrons in turn (for example, from B to Ne), one gets therefore only \( p \)-terms. This is in accordance with experience. In particular, in the case of three electrons in an eight-electron shell, there results the well-known alkali-like doublet spectrum, and no \( s \)-term appears which has the same principal quantum number as that of the ground state. The case of four electrons is related to the arc-discharge spectrum of carbon, not analyzed yet, and to that of lead, partly analyzed so far. Pauli points out the necessity of discriminating the three ways of combination of
two electrons with the set of \( n, k_1 \) and \( k_2 \), denoted by \( n_{k_1}k_2 \), for the case of \( k_1 = 2 \) and \( k_2 = 1 \) or 2. Namely, they are the combinations of two equivalent electrons with \( n_{21} \), of one electron with \( n_{22} \) and one electron with \( n_{21} \), and of two electrons with \( n_{22} \). He shows, through the calculation of \( m_j \) for each combination, that there exist five \( p \)-terms, of which two are with \( j = 2 \), two are with \( j = 0 \), and the last one is with \( j = 1 \). He gets also the \( g \)-factor of these terms by calculating the value of \( m_2 \). For the discussion of the cases of more than four, up to eight electrons, it suffices to apply the reciprocity law.

The "general rule" introduced here by Pauli is the one which was later given the name the exclusion principle. In Bohr's theory of atomic structure, it is assumed that when one bound electron is added to an atom, the quantum numbers of electrons already bound in the atom do not alter, and this group of electrons becomes the atomic core of the newly formed atom (the principle of invariance of quantum numbers, or the principle of formation of atoms in Bohr's saying). This faces with the contradictions, as is mentioned above, in the case of the multiplet terms of the alkaline earths. Thus Pauli successfully solves the difficulties in the theory of atomic structure as well as arrives at "general rule", by attributing the doublet structure of the alkali spectrum to the "double-ness" of the electron, and by connecting Stoner's configuration of atomic electrons with the Paschen-Back effect in the anomalous Zeeman effect with the use of the adiabatic theorem.

By finding a new attribute of the "double-ness" of the electron, Bohr's "principle of formation of atoms" is brought up to Pauli's "general rule," and the related phenomena become understood in a more unified way. Pauli concludes his paper by saying that the problem of founding the general rule "might be successfully effected first of all according to a further deepening of the fundamental principles of the quantum theory". This task is later accomplished by finding in quantum mechanics the symmetry and the anti-symmetry of the wave function for the interchange of two identical particles.

§3. The introduction of spin

About ten months after the publication of Pauli's paper mentioned in the preceding subsection, Uhlenbeck and Goudsmit interpret the attribute of the double-ness of electron found by Pauli, by means of the substantialistic model of the self-rotation of electron, in their paper "Replacement of the Hypothesis of Non-mechanical Compulsion by a Demand about the inner Behaviour of each
individual Electron". They say that "Pauli has already struck a new way". They consider, however, that "Pauli does not tie himself down to any idea of a model," and "to us another way seems still open," and thus take notice of the four quantum numbers of each individual electron. Uhlenbeck-Goudsmit think that "the four quantum numbers have lost their original Landé's meaning," and "each electron is now given also four degrees of freedom together with the four quantum numbers".

Saying that "one can then give the following meaning to the quantum numbers," they assign an "eigen-rotation" to Landé's $R$ (the angular momentum of the atomic body). The principal, azimuthal and magnetic quantum numbers, $n$, $k$ and $m$ respectively, remain to have their old meanings. Namely, Uhlenbeck-Goudsmit see "the classically unexplainable double-ness of electron," which is thought by Pauli as giving rise to "the not yet understood property which Landé ascribed to the atomic core," again from the side of angular momentum, introducing the model of "the eigen-rotation" of electron.

They say that "further quantitative accomplishment of this idea will well depend strongly on the choice of the model of electron", and "to come to agreement with experience, one must therefore put the following demands on this model". Firstly, the ratio of the magnetic moment to the angular momentum for the eigen-rotation must be twice as large as that for the orbital motion. In this relation, they cite Abraham's result that the ratio under consideration is twice the usual value, if a self-rotating sphere which is of a radius equal to the classical electron radius and has a unit electric charge uniformly distributed on its surface is taken as a model of electron. They point out, at the same time, that when its angular momentum is quantized, the rotating velocity on the surface becomes larger than the light velocity. Secondly, various orientations of $R$ relative to the orbital plane must be able to provide the explanation of the relativity doublet.

Uhlenbeck and Goudsmit treat this problem two months later in their paper "Spinning Electrons and the Structure of Spectra". They use here the expression "spin" in place of "eigen-rotation" in their preceding paper. By the way, the expression "spinning electron" was already used by Compton in 1921 in his treatment of the magnetism of matter.

---

Compton discusses in the paper "The Magnetic Electron"\(^{75}\) the origin of the magnetism of matter, making reference to Barnet's and Stewart's papers\(^{76}\) which we have taken in Sec. 1, §3 of the present chapter. Compton says that if the magnetism is supposed to result from the revolution of electrons in orbit, "the diamagnetism can be more than a transient phenomenon," because "when an external field is applied the induced change in angular momentum will put the electrons in an unstable condition".

Compton then examines, in regard to Person's work,\(^{77}\) the possibility that the electron possesses a magnetic moment as a result of its spinning motion. Considering that some non-electric bond would be necessary for the hydrogen molecule and others, person supposed about the electron that "the unit negative charge is distributed continuously around a ring which rotates around on its axis," and discussed molecular bindings by means of the magnetic forces among electrons. Compton calls the self-rotation of the electron spinning motion. He says that if the electron possesses an angular momentum of \(\hbar/2\pi\) for its spinning motion, the size of the electron should be greater than \(10^{-10}\) cm for the peripheral velocity of the electron not to approach the light velocity. He considers that this is favorable in view of the scattering of X-rays and \(\gamma\)-rays. He says also that if the orbital motion of electrons is the cause of magnetism, a change in the intensity of X-ray reflection from a crystal of magnetite would be expected when the crystal is strongly magnetized, on account of the change in the planes of the revolution of the electrons, but any measurable effect is not detected.\(^{78}\)

Compton appreciates Allen's explanation\(^{79}\) of the rotation of the plane of polarization by optically active substance, on the assumption of a magnetic and sizable electron, as a strong evidence that electron possesses a magnetic moment. He discusses a phenomenon, first noticed by C. T. R. Wilson and brought to Compton's attention by Mr. Shimizu,\(^{80}\) as a "direct evidence" that free electrons possess magnetic polarity. It is observed in photographs taken by Wilson of the tracks of secondary cathode rays and beta particles.

\(^{75}\)A. H. Compton, Journ. of Franklin Ins. 192 (1921), 145.
\(^{76}\)Cf. Refs. 23) and 24) of the present chapter.
\(^{77}\)A. L. Person, Smithsonian Misc. Collections, 1915.
\(^{79}\)H. S. Allen, Phil. Mag. 40 (1920), 426.
\(^{80}\)Mr. Shimizu referred to by Compton is probably T. Shimizu who improved in 1921 at Rutherford's laboratory the Wilson chamber to work repeatedly.
Paths in the form of almost perfect spirals are seen there. Of this phenomenon Compton gives the interpretation that, when a “magnetic electron” is placed in a homogeneous paramagnetic medium, a part of the medium surrounding the electron will be slightly magnetized, and the induced magnetization will give rise to a magnetic force at the electron to make its path spiral. Compton points out also that magnetic electrons will remove the magneto-mechanical anomaly, and gives qualitative discussions of the magnetic moment of a spinning electron, by supposing generally that the angular momentum of the spinning motion is $\hbar/2\pi$.

Now, Uhlenbeck and Goudsmit consider in their second paper, having mentioned that they were “without being aware of Compton’s suggestion,” the effect of the spin on the orbital motion of the electron. Because the electron has a magnetic moment due to the spin, “the electron will be acted on by a couple just as if it were placed at rest in a magnetic field of magnitude equal to the vector product of the nuclear electric field and the velocity of the electron relative to the nucleus divided by the velocity of light”. This is the original idea of the so-called spin-orbit interaction in modern physics. The couple will cause a slow precession of the spin axis, giving rise to a set of states which differ in the orientation of the spin axis relative to the orbital plane. “If the spin corresponds to a one-quantum rotation there will be in general two such states.”

These considerations bring about “an essential modification of the explanation hitherto given of the fine structure of the hydrogen-like spectra”. The scheme given by them of the energy levels of the hydrogen atom consists of $S_{1/2}, P_{3/2}, D_{5/2}, \ldots$ in the case of no spin, where $1/2, 3/2, 5/2, \ldots$ correspond to the values of Landé’s orbital angular momentum $K = k - 1/2$. The summing up of the spin angular momentum $R = 1/2$ to this results in the total angular momentum of the electron $J(1/2 \rightarrow 1, 3/2 \rightarrow 1$ and $2, 5/2 \rightarrow 2$ and $3, \ldots)$. At the same time, there results a change in the energy of the electron on account of the spin magnetic moment, so that two energy levels of a given value of $J$ turn out to lie adjacent to each other near the energy level of the value of $k$ equal to $J$ in the old theory, giving a new explanation to the fine structure of spectral terms.

Uhlenbeck and Goudsmit say that the new theory explains the occurrence of certain components in the fine structure of the hydrogen spectrum, which according to the old scheme would correspond to transitions in which $K$ remains unchanged. According to the new scheme, in the transitions in question...
Theory of Atomic Structure and Spin of Electron

$K$ changes by one unit and only $J$ remains unchanged, quite in conformity with the correspondence principle. They mention that “the assumption of the spinning electron leads to a new insight into the remarkable analogy between the multiplet structure of the optical spectra and the structure of the X-ray spectra,” and that “it may be possible to account for the important results arrived at by Pauli, without having to assume non-mechanical ‘duality’ in the binding of the electrons”.

They stress that the difficulties in the anomalous Zeeman effect “disappear at once”. This is because the coefficient of magnetic moment of the spin is different from that of the orbital motion, and the spin axis precesses with a frequency different from the Larmor frequency. They interpret the Paschen–Back effect by considering that, when the magnetic field is strong, its influence on the precession of the spin axis becomes comparable with that due to the orbital motion.

The assumption of the anomalous coefficient of magnetic moment of the spin, however, “leads to widths of the spin doublets just twice as large as those required by observation”. This difficulty is solved by Thomas\textsuperscript{81} about two months later. Thomas says that the spin axis is required to precess around the magnetic field $H$ with the angular frequency $\omega = eH/mc$. The magnetic field due to the nucleus is given by $H = E \times (u/c)$, where $E$ and $u$ denote the electric field due to the nucleus and the velocity of the electron, respectively. The value of $\omega$ obtained from this magnetic field gives the double value for the width of the doublet term due to the spin, as is mentioned above.

The magnetic field $H$ given above is obtained by means of the Lorentz transformation from the rest system of coordinates of the nucleus to the system of coordinates moving with the velocity $u$ relative to the nucleus. Thomas points out that “there is, however, an error in the above reasoning”. Because the electron has the acceleration $f = -eE/m$, its velocity after a time duration $dt$ becomes $u + fdt$. He considers that the precession to be obtained is to rotate the spin axis in the rest systems of coordinates of the electron, $S$, at the time $t$, to that in the rest system of coordinates of the electron, $S'$, at the time $t + dt$. The systems $S$ and $S'$ are related approximately through the Lorentz transformation with the velocity $f dt$ and the rotation $u \times f/2c^2$. The resulting value of $\omega$ is therefore smaller by $u \times f/2c^2$ than its value shown

\textsuperscript{81}L. H. Thomas, Nature 117 (1926), 514.
above, so that it is just half the latter value, in agreement with the required value.\(^\text{82})\)

The energy, \(\pm \hbar \omega / 2\pi\) due to the interaction between the spin and the orbital motion obtained by relativistic kinematics is proportional to \((Z - s)^4\) for a Coulomb field \(E = (Z - s)e r / r^3\) with screening constant \(s\), because \(r \times u\) is expressed in terms of the angular momentum quantum number and the average value of \(1 / r^3\) is proportional to \((Z - s)^3\), on account of the quantization of the orbital motion. In Landé's model,\(^\text{83})\) because \(E\) was taken to be the Coulomb field \(-e r / r^3\) at the nucleus due to the electron, the interaction energy between the magnetic moment assumed for the atomic body and the orbital motion of the electron was proportional to \((Z - s)^3\).

As we have analyzed in Sec. 2, §3, Chap. 3 of the present volume, Sommerfeld's relativistic formula for the energy level possesses a limitation that is characteristic of the lawfulness in a substantialistic stage. The second term of this formula, that is the fine structure term, is now replaced, on account of the discovery of a substantialistic attribute of the electron named the spin, by the energy of the substantialistic spin-orbit interaction, which is derived from the quantum condition imposed on the non-relativistic orbital motion and from relativistic kinematics. The revolution of the perihelion of the orbit in the relativistic Kepler motion is replaced by the self-rotation of the electron in the non-relativistic orbital motion, and from relativistic kinematics. The revolution of the perihelion of the orbit in the relativistic Kepler motion is replaced by the self-rotation of the electron in the non-relativistic orbital motion. The essentialistic unification of the quantum theory and the theory of relativity is achieved later by the Dirac equation of the electron.

By the way, in the issue of Nature next to that in which Thomas' paper is published, Kronig discusses the conceptual difficulties involved in the idea of spinning electron. In his paper "Spinning Electrons and the Structure of Spectra,"\(^\text{84})\) Kronig takes up first the difficulty of the internal velocities of the electron, which is pointed out also in Uhlenbeck-Goudsmit's paper. He states next that the elementary unit of magnetic moment, the Bohr magneton, is derived from considerations of the orbital motion of electrons with velocities

\(^{82})\) J. Frenkel, being inspired by Thomas' paper, treated this problem with a relativistic four dimensional formulation [Zeit. f. Phys. 37 (1926), 243].

\(^{83})\) Cf. Sec. 2, §4, p. 211 of the present chapter.

\(^{84})\) R. de Kronig, Nature 117 (1926), 550.
much smaller than that of light, in which the internal structure of these electrons does not enter at all, and "it is hard to see, then, why this elementary unit should also be characteristic of the internal motion of the electrons in spite of the high velocities involved".

Furthermore, Kronig considers that, when an electron forms part of the nuclear structure, the nucleus will have a magnetic moment of the order of a Bohr magneton, unless the magnetic moments of all the nuclear electrons just happened to cancel. He mentions that "for such an additional moment of the nucleus there is no place in the theory of the Zeeman effect, and the probability that in all atomic nuclei the magnetic moments of the electrons neutralise seems a priori to be very small". He concludes his paper by saying that "the new hypothesis, therefore, appears to effect the removal of the family ghost from the basement to the sub-basement, instead of expelling it definitely from the house".

The second difficulty indicated by Kronig appears from the characteristic mentioned above of the theory of the spin-orbit interaction that remains in a substantialistic stage. Needless to say, the third difficulty pointed out by Kronig is solved later, when a substance named the neutron is discovered, so that the electron is 'expelled definitely' from the nucleus. The first difficulty mentioned by Kronig is connected with the problem of the size of the electron. At that time, it was used to suppose for the electron the classical radius $e^2/mc^2 \approx 3 \times 10^{-13}$ cm in order to avoid the self-electrostatic energy of the electron becoming infinitely large. The first difficulty in question is solved later by Dirac's relativistic wave equation of the electron, the solution of which includes of itself the freedom of the spin without any assumption on the size of the electron.

However, as soon as the interaction between the wave field of electron and the electromagnetic wave field is treated by means of the quantum theory of fields, there arises anew the problem of the divergence of the self-energy of the electron, etc. Though the solution to this problem is given for the time being by the renormalization theory due to Feynman, Schwinger and Tomonaga, in which finite differences between infinite quantities are assumed to have actual meaning in the quantum electrodynamics, following Sakata's theory of C-meson which mediates the cohesive force for the electron, the problem of the structure of the electron still remains, together with that of other elementary particles, to be one of the fundamental problems in the present-day elementary particle physics. In the first difficulty mentioned
in Kronig's paper is thus seen the germ of such a fundamental theoretical problem.\textsuperscript{85)}

\textsuperscript{85)} Of this problem, Taketani presented at a meeting of the physical society of Japan held in 1943 some analyses of various aspects of the interaction between the field and matter. From this, as a starting point, discussions spread and Sakata's theory of C-meson was afterward presented. Taketani's analyses were published in his paper "Conflict between Field and Matter" [\textit{Kagaku (Science)}], Nov. Issue, 1946; put in M. Taketani, "Problems in Dialectics", (Vol. I of "Collected Works").
This page is intentionally left blank
Chapter 5

The Interconnection of Wave- and Particle-Natures
Barkla-White's experiment (1917)
X-ray absorption smaller than that in classical theory.

Ishino's experiment (1917)
γ-ray absorption smaller than that in classical theory.

1917
Compton (1919)
Model of an electron with extension to reduce X-ray scattering by interference.

1917
Compton (1919)
Attempt to derive the quantum condition by an inneratomic elastic wave.

1917
Bohr (1923)
Doubt of the light quantum theory and of the energy-momentum conservation in an elementary process.

1917
Duane (1923)
Derivation of the Bragg formula by the light quantum theory and the quantum condition for momentum change of a lattice.

1917
Brillouin (1919)
Application of transition probability to the dispersion theory of light.

1917
Schrodinger (1922)
Derivation of Doppler effect from the light quantum theory and frequency relation.

1917
Debye (1921)
Application of transition probability to the dispersion theory of light.

1917
Ladenberg (1921)
Calculation based on the light quantum theory of Compton effect, and its experimental verification.

1917
Debye (1923)
Calculation as above.

1917
Ladenberg (1923)
Theory of cavity radiation by means of light quanta with a mass.

1917
De Broglie (1923)
Derivation of the quantum condition and prediction of electron diffraction by the idea of accompanying phase wave.

1917
De Broglie (1923)
Derivation of the light quantum theory and the Bragg formula from the quantum condition for periodical motion in space.

1917
Brillouin (1919)
Calculation based on the light quantum theory of Compton effect, and its experimental verification.

1917
Schrödinger (1923)
Derivation of Doppler effect from the light quantum theory and frequency relation.

1917
Debye (1923)
Application of transition probability to the dispersion theory of light.

1917
Ladenberg (1921)
Calculation based on the light quantum theory of Compton effect, and its experimental verification.

1917
Debye (1923)
Calculation as above.
3, §2
Borh-Kramers-Slater (1924)
Hypothesis of statistical conservation of energy-momentum. Theory of classical interaction between virtual oscillators and light.

3, §3
Bothe-Geiger (1925)
Negation of the statistical conservation of energy-momentum by a coincidence experiment.
Compton-Simon (1925)
The same by a cloud chamber experiment.

3, §1
Slater (1924)
Idea of virtual radiation field of frequencies corresponding to transitions.

4, §1
Kramers (1924)
Extension of Ladenberg's dispersion formula by the application of induced radiation.

4, §2
Kramers-Heisenberg (1925)
Generalization of Kramer's dispersion formula in accordance with Smekal's result. Dipole moment with components related to transitions.

3, §1
Slater (1924)
Theory of dispersion and scattering of light on the basis of the light quantum theory.

4, §2
Smekal (1923)
Theory of dispersion and scattering of light on the basis of the light quantum theory.

2, §4
Bose (1924)
Attempt to derive Planck's formula by the light quantum theory.

Einstein (1924)
Quantum theory of gases obeying B-E statistics. Decrease in viscosity at low temperature due to the wave nature of molecules.

2, §4
Ellassel (1925)
Interpretation of data on scattering and reflection of electrons with the wave theory.

To Heisenberg's matrix mechanics
To Schrödinger's wave mechanics

Note: The numbers above a frame indicate Sec. and Subsec.
Chapter 5

The Interconnection of Wave- and Particle-Natures

Theme in Chapter 5

Although the calculation of the Compton effect on the basis of the theory of light quantum is well known through text books in modern physics, it is not so well known that Compton attempted to explain the result of X-ray scattering by means of the effect of interference, but confronted with a serious difficulty, and thus turned to the theory of light quantum.

About the time of the discovery of the Compton effect, Schrödinger derived, by making use of the conservation law of energy and momentum, from the standpoint of the theory of light quantum, the formula for the Doppler effect, and Duane, by applying the quantum condition and also from the viewpoint of the theory of light quantum, the Bragg formula for X-ray scattering. It is interesting to see the process by which they came to apply the theory of light theory to wave phenomena.

It should be noted especially that de Broglie’s wave theory actually resulted from the theory of light quantum. As we shall see in the text, de Broglie’s attempt to understand the particle-nature of light quantum by means of the rest mass of light quantum brought forth, rather, the concept of the wave-nature of material particles.

Bohr, who was in doubt about the theory of light quantum, proposed in collaboration with Kramers and Slater the law of statistical conservation of energy and momentum and treated the dispersion of light as a typical wave phenomenon. They proposed the law of statistical conservation of energy and momentum and treated the dispersion of light as a typical wave phenomenon. Their proposal of the law of statistical conservation of energy and momentum was, however, negated immediately by Compton and Simon’s experiment with the use of a Wilson chamber.

It is scarcely told, so far, in textbooks on the history of science that Kramers and Heisenberg, in generalizing Kramers’ wave-theoretical formula for the dispersion of light to include the case of incoherent scattering of light, took account of Smekal’s study of the scattering of light on the basis of the theory of light quantum.
1. Grasping of Wave Phenomena by Means of the Theory of Light Quantum

§1. The theory of light quantum and the Doppler effect

As we have already seen in Vol. I and in Chap. 1 of the present volume, the theory of light quantum presented in 1905 by Einstein gave the Planck distribution formula at short wave lengths, that is the Wien distribution formula. As has been described in Sec. 2, Chap. 3 of the present volume, Einstein introduced in 1916 the concept of transition probability for emission and absorption of radiation by each individual atom, and thereby derived the Planck distribution formula.

In Einstein's theory of light, a ray of light of frequency \( \nu \) was considered to consist of light quanta of energy \( h\nu \), and momentum \( h\nu/c \). Needless to say, the particle-nature of light brought in by the theory of light quantum was in conflict with the wave theory of light so seriously that there came Lorentz's paper in 1910 which we have mentioned in Sec. 3, §3, Chap. 1 of the present volume. However, Einstein's theory of radiation in 1916, in which the Maxwell distribution of velocities of molecules in equilibrium with radiation as well as the Planck distribution formula was derived, opened the way to grasp typical wave phenomena of light on the basis of the theory of light quantum.

In a paper in 1922 “The Doppler Principle and Bohr's Frequency Relation”,¹ Schrödinger discusses the Doppler effect from the stand point of the theory of light quantum. He starts from critical comments on Försterling's paper "Bohr's Atomic Model and the Theory of Relativity".² Försterling attempted to derive the formula of the Doppler effect by connecting Bohr’s frequency relation with the theory of relativity. Schrödinger considers that Försterling’s result is “disappointing”. Försterling supposed that the frequency \( \nu \) of light emitted from an atom in its rest state of energy \( E \) was related to the frequency of light \( \nu' \) emitted from the same atom in its moving state of energy \( E' \), through the formula

\[
\frac{\nu'}{E'} = \frac{\nu}{E}
\]

"on account of Bohr’s frequency relation", together with the assumption of the invariant Planck constant.

¹E. Schrödinger, Phys. Zeit. 23 (1922), 301.
²K. Försterling, Zeit. f. Phys. 3 (1920), 404.
From the formula $E' = E/\sqrt{1-u^2/c^2}$ for the atom moving with a velocity $u$, Försterling got

$$\nu' = \frac{\nu}{\sqrt{1-u^2/c^2}}. \quad (2)$$

He said that this was in agreement with the formula of the transverse Doppler effect, namely, the Doppler effect for the light coming from a source moving transverse to the line of sight.\(^3\)

Försterling's formula (2) does not explain the longitudinal Doppler effect, namely, that for the light coming from a source moving in the direction of the line of sight. Furthermore, in Försterling's calculation the velocity of the source is supposed not to change, in spite of his consideration of an atom as the source. Schrödinger criticizes Försterling's idea in this way and says that "according to Einstein's foundation of the theory of radiation, the emitted light quantum $h\nu$ always carries, in any system of coordinates, the momentum $h\nu/c$", citing Einstein's paper\(^4\) in 1917 on the transition probability of an atom. In Försterling's paper the "recoil" of the atom due to the momentum of the light quantum is not taken into account.

Schrödinger takes the case of the emission of light from a molecule, and denotes the rest masses of the molecule before and after the emission of light by $E_1/c^2$ and $E_2/c^2$, respectively. $E_1$ is the energy of the molecule in a stationary state in the rest system of coordinates before the emission, while $E_2$ is that after the emission. In the observational system, which is "called simply the spectrometer", the molecule is assumed to have the velocities $u_1$ and $u_2$ before and after the emission, respectively. Considering first the case of the velocity of the molecule being in the same direction as the momentum of the light quantum, for simplicity, the energies of the molecule before and after the emission are given by $E_1/\sqrt{1-u_1^2/c^2}$ and $E_2/\sqrt{1-u_2^2/c^2}$ respectively with regard to the spectrometer.

"Bohr's frequency relation becomes

$$h\nu = \frac{E_1}{\sqrt{1-u_1^2/c^2}} - \frac{E_2}{\sqrt{1-u_2^2/c^2}}. \quad (3)$$

\(^3\)Försterling discussed the red shift in the general theory of relativity, by supposing that an atom with rest mass $E/c^2$ would have the energy $E' = E + (E/c^2)\Phi$ at a point of the gravitational potential $\Phi$, giving the relation $\nu' = \nu(1 + \Phi/c^2)$.

\(^4\)A. Einstein, the same as Ref. 41) in Sec. 3, §2, Chap. 3 of the present volume.
Furthermore, there is the balance of the momentum with regard to the spectrometer

\[ \frac{E_1 u_1}{c^2 \sqrt{1 - u_1^2/c^2}} = \frac{E_2 u_2}{c^2 \sqrt{1 - u_2^2/c^2}} + \frac{h\nu}{c}. \]  

(4)

"Bohr’s frequency relation” in Schrödinger’s saying, Eq. (3), represents the law of conservation of energy seen from the spectrometer.

"E₁ and E₂ take the values naturally determined according to the character of the quantum transition”, says Schrödinger. Thus, when \( u_1 \) is given, \( \nu \) and \( u_2 \) are obtained from Eqs. (3) and (4). Putting

\[ \nu^* = \frac{E_1^2 - E_2^2}{2h\sqrt{E_1 E_2}}, \]

there results the formula

\[ \nu^* = \nu \left\{ \frac{c - u_1}{\sqrt{c^2 - u_1^2}} \cdot \frac{c - u_2}{\sqrt{c^2 - u_2^2}} \right\}^{\frac{1}{2}}. \]

(6)

Schrödinger considers that this is to be compared with the usual formula for the Doppler effect

\[ \nu^* = \nu \frac{c - u}{\sqrt{c^2 - u^2}}, \]

(7)

in which \( \nu^* \) denotes the frequency in the rest system of the source, and \( \nu \) the frequency in the case of the source moving with the velocity \( u \) relative to the observer.

Schrödinger considers that \( \nu^* \) defined in Eq. (5) “plays the role of the frequency at rest”. Equation (6) includes a factor which is equal to the geometrical mean of the values of the factor in Eq. (7) for the \( u \)’s before and after the emission, \( u_1 \) and \( u_2 \) respectively. He says that this is “a natural generalization of the relativistic Doppler principle”, and “it is self-evident that different velocities appear before and after the emission” in this factor. The change in the velocity of the molecule due to emission of the light quantum is “very small compared with the thermal velocity”. If the difference between \( u_1 \) and \( u_2 \) is neglected, Eq. (6) becomes identical with Eq. (7) for the source moving with a constant velocity. He says the “the change in the velocity of a molecule having the mean thermal velocity amounts only to one thousandth,
even in the case of a favorable condition" that an ultraviolet ray of wavelength 200 Å is absorbed in a gas of low temperature.\textsuperscript{5}

For the case where the molecule moves with velocities of magnitudes $u_1$ and $u_2$ before and after the emission, as in the above case, but in directions different from that of the momentum of the light quantum by angles seen from the spectrometer $\theta_1$ and $\theta_2$ before and after the emission, respectively, Schrödinger shows that Eq. (6) is to be replaced by

\begin{equation}
\nu^* = \nu \left\{ \frac{c - u_1 \cos \theta_1}{\sqrt{c^2 - u_1^2}} \cdot \frac{c - u_2 \cos \theta_2}{\sqrt{c^2 - u_2^2}} \right\}^{\frac{1}{2}}, \tag{6'}
\end{equation}

where $\nu^*$ has the same meaning as above. This equation coincides with the formula of the relativistic Doppler effect for a source moving with a constant velocity too, if the difference between $u_1$ and $u_2$ is neglected.

In this way, Schrödinger derived the formula for the Doppler effect, which had been understood to be a typical wave phenomenon, by applying the theory of light quantum to the elementary process of the emission of light by a molecule. In the following year, 1923, a most elementary process, the scattering of X-ray by an electron, was treated by Compton and Debye independently, from the standpoint of the theory of light quantum, and their result was confirmed experimentally by Compton.

\section*{§2. The theory of light quantum and the Compton effect}

In a paper received on Dec. 13, 1922 and published in May, 1923, "A Quantum Theory of the Scattering of X-rays by Light Elements",\textsuperscript{6} Compton gave discussions of the Compton effect, so called nowadays, and showed evidences for it. Experiments on the scattering of X-rays by atoms had been carried out previously by Barkla and others, and the results had been considered to fit to the Thomson scattering formula, which was calculated for the scattering of electromagnetic waves by an free electron with the use of the classical

\textsuperscript{5}R. Mössbauer [Zeit. f. Phys. 151 (1959), 125] made the change in velocity vanish, by using as the emitting or absorbing body a crystal of Os at such an extremely low temperature that no oscillation was excited in the crystal, and thereby putting the change in momentum accompanying the emission and absorption of the $\gamma$-ray (129 keV) from Ir$^{191}_{191}$ due to the $\beta$-decay of O$_{191}$ upon the whole crystal.

\textsuperscript{6}A. H. Compton, Phys. Rev. 21 (1923), 483.
theory.\(^7\) According to the Thomson formula, the intensity of scattered X-rays is independent of their wavelength, and the wavelength of X-rays scattered by a thin layer of matter is identical on both the side of the layer. As X-rays of shorter wavelengths and γ-rays came into use, experimental results began to show considerable differences from the Thomson formula.

Barkla and White\(^8\) showed that for light elements the absorption coefficients of X-rays of wavelengths shorter than \(2 \times 10^{-9}\) cm were smaller than the values calculated on the assumption that the absorption results only from Thomson scattering. Ishino\(^9\) showed that for light elements of atomic weights less than 32, the absorption coefficients of γ-rays of high frequencies were considerably smaller than those of the usual X-rays, and the absorption coefficients per unit mass were several times smaller than the values expected from the Thomson formula.

In a 1919 paper “The Size and Shape of the Electron”,\(^10\) Compton accordingly attempts to suppose a size for the electron. He says that “if the electron is considered to have a radius comparable with the wavelength of the incident radiation, a qualitative explanation of the phenomenon of low scattering for short wavelengths is obvious”. This is because the hypothesis is effective “to make an appreciable phase difference between the rays scattered by different parts of the electron”. Compton considers as the form of electron a rigid uniform spherical shell of electricity incapable of rotation, a flexible spherical shell of electricity capable of independent displacement of each part and a thin flexible ring of electricity. By comparing the results of calculation with Ishino’s experimental data, Compton deduces as the radius of the electron the values 1.7, 2.1 and 2.7 in units of \(10^{-10}\) cm, respectively, for the cases considered. If the value \(2 \times 10^{-10}\) cm is taken for the radius of the electron, the asymmetry observed by Florance\(^11\) of the intensity of γ-rays scattered at the forward and backward directions can be explained.

The result of an experiment made by Compton\(^12\) in 1922, however, shows that the radius of the electron, \(a\), should increase with the wavelength \(λ\) of X-rays. It is a natural circumstance that one has the relation \(a \sim λ\), since

---

7) Cf. Sec. 6, §1, Chap. 2, Vol. I.
8) C. G. Barkla and M. P. White, Phil. Mag. 33 (1917), 277.
9) M. Ishino, Phil. Mag. 33 (1917), 129.
the intensity of the scattered radiation is made weak by means of interference effect. Further, a much more serious difficulty had been found previously in 1921 by Compton himself.\textsuperscript{13} It is the fact that only a small part, if any, of the secondary X-radiation is of the same wavelength as the primary, from a spectroscopic examination of the secondary X-rays from graphite.

Compton says, in his 1923 paper on the Compton effect cited above, that “such a change in wavelength is directly contrary to Thomson’s theory”. For, Thomson’s theory demands that “the scattering electrons, radiating as they do because of their forced vibrations when traversed by a primary X-ray, shall give rise to radiation of exactly the same frequency as that of the radiation falling upon them”. “Nor does any modification of the theory such as the hypothesis of a large electron suggest a way out of the difficulty.” Compton concludes, thus, that “this failure makes it appear improbable that a satisfactory explanation of the scattering of X-rays can be reached on the basis of the classical electrodynamics”.

He turns then to treat the scattering of X-rays by an free electron from the standpoint of the theory of light quantum. In this case, “the scattering electron will recoil with a momentum equal to the change in momentum of the X-ray”. The energy of the scattered ray will be less than that of the incident one by the kinetic energy of the recoil of the electron. Compton thinks that “since the scattered ray must be a complete quantum, the frequency will be reduced in the same ratio as is the energy”. From these considerations, it results that when a X-ray of wavelength $\lambda_0$ is incident upon a free electron at rest, the wavelength $\lambda_\theta$ of the X-ray scattered at the angle $\theta$ is given by

$$\lambda_\theta = \lambda_0 + 2 \left( \frac{h}{mc} \right) \sin^2 \theta, \quad (8)$$

where $m$ denotes the rest mass of the electron. $h/mc$ is the quantity which later becomes known as the Compton wave length of the particle of mass $m$.

Compton points out that, if the incident wave of wavelength $\lambda_0$ is supposed to be scattered by an electron moving with an “effective velocity” $\bar{\beta}c$ in the direction of the incident wave, the Doppler effect in the classical theory will give the same change in wavelength as is given in Eq. (8), with $\bar{\beta}$ defined by

$$\bar{\beta} = \frac{\alpha}{(1 + \alpha)^{\frac{1}{2}}}, \quad \alpha = \frac{(h/mc)}{\lambda_0}. \quad (9)$$

\textsuperscript{13}A. H. Compton, Phil. Mag. 41 (1921), 749; Phys. Rev. 18 (1921), 96.
According to Eq. (8), the change in wavelength, $\lambda_{\theta} - \lambda_0$, is independent of the initial wavelength $\lambda_0$. It is already known to Compton\textsuperscript{14}) that the increase in wavelength of the secondary X-rays at 90° is roughly 0.03 Å for primary X-rays of wavelength 0.7 ~ 0.025 Å. Furthermore, from a new graphical study of the result of his experiments previously performed on the scattering of the characteristic K-lines from molybdenum by graphite at 90°, Compton finds that the spectrum of the scattered X-rays is closely similar to that of the primary X-rays in form, but is shifted from the latter spectrum in wavelength by 0.022 Å, in good agreement with the theoretical value 0.024 Å from Eq. (8).\textsuperscript{15}

Now, in order to investigate the spatial distribution of the energy of the scattered radiation, Compton assumes that the secondary light quantum is emitted from the electron moving with the effective velocity, and the probability of emission is proportional to the angular distribution of the intensity of the Thomson scattering in the rest coordinate system of the electron, that is, $(1 + \cos^2 \theta')$ where $\theta'$ is the scattering angle in the system mentioned. He is not successful in showing that even the spatial distribution of the intensity in the actual scattering could be treated with the effective velocity in the framework of the classical theory. Nevertheless, he assumes this hypothesis to calculate the intensity of the scattered beam, $I_\theta$ at the angle $\theta$ in the laboratory coordinate system. Let $I$ be the intensity of the incident beam of X-rays, $N$ the number of the scattering electrons per unit volume, and $n$ the number of the light quanta of the scattered beam. Compton expresses $I_\theta$ in terms of $N$, $\theta$ and $n$, making use of the Lorentz transformation from $\theta'$ to $\theta$ and the change of wavelength at $\theta$ given by Eq. (8). Because at $\theta = 0$ there is no recoil of the electron, he considers that $I_\theta$ “should be that calculated on the basis of the classical theory”. From this relation $n$ is expressed by $I$, so that $I_\theta$ is determined by $N$, $I$ and $\theta$.

As a result, the absorption coefficient of X-rays due to scattering is given by

$$\sigma = \frac{\sigma_0}{(1 + 2\alpha)}; \quad \sigma_0 = \left(\frac{8\pi}{3}\right) \frac{Ne^4}{m^2c^4}.$$  \hspace{1cm} (10)
\( \sigma_0 \) is the scattering coefficient due to the classical theory. Compton derives, in this way, the angular distribution of the intensity of scattered X-rays different from the Thomson scattering formula, by choosing on the basis of the theory of light quantum situations to which some classical relations seem to be applicable.

The absorption coefficient per unit mass \( \mu/\rho \) has been known to be given by \( \mu/\rho = \kappa \lambda^3 + \sigma/\rho \) for a wide range of wavelength, where \( \rho \) denotes the density and \( \kappa \) is a constant. Compton determines the value of \( \kappa \) for carbon from the long wavelength part of Hewlett's observational results of \( \mu/\rho \), and subtracts the values of \( \kappa \lambda^3 \) with \( \kappa \) determined from Hewlett's \( \mu/\rho \) for various \( \lambda \), to deduce the value of \( \sigma/\rho \). Compton shows that the value of \( \sigma/\rho \) thus deduced is in good agreement with that calculated from Eq. (10) for wavelengths less than 0.5 Å. In Ishino's paper cited above, the observed values are given about the true absorption coefficient per unit mass, namely, the ratio of the energy obtained by subtracting the energy going out as the scattered rays from the energy removed from the incident beam, for the scattering of \( \gamma \)-rays from RaC by Al and Fe. The values of the true absorption coefficient calculated by Compton on the assumption of the effective velocity are 0.040 and 0.038 for Al and Fe, respectively, "which do not differ seriously from the experimental values" 0.045 and 0.042 respectively.

Compton mentions that "electrons which recoil in the process of the scattering of ordinary X-rays have not been observed". This is "probably because their number and velocity are usually small compared with the number and velocity of the photoelectrons ejected as a result of the characteristic fluorescent absorption". He says also that "there is good reason for believing that most of the secondary \( \beta \)-rays excited in light elements by the action of \( \gamma \)-rays are such recoil electrons".

By the way, about the Thomson scattering formula Debye considers, in a paper received on Mar. 14 and published in Apr. in 1923 "Scattering of X-rays and Quantum Theory", that "all the experimental results for short waves

---

\(^{16}\) \( \sigma_0/N \) is the quantity that is called the Thomson cross section nowadays. According to quantum electrodynamics in which the Dirac relativistic equation is used for the electron, \( \sigma/\sigma_0 = 1 - 2\alpha + (26/3)\alpha^2 + \cdots \) for \( \alpha \ll 1 \), and \( \sigma/\sigma_0 \approx (3/8\alpha)(\log 2\alpha + 1/2) \) for \( \alpha \gg 1 \) [O. Klein and Y. Nishina, Zeit. f. Phys. 52 (1929), 853; the Klein-Nishina scattering formula; cf. Sec. 3, §7, Chap. 3, Vol. III].

\(^{17}\) C. W. Hewlett, Phys. Rev. 17 (1921), 284.

stand in contradiction”. He says, citing Compton’s paper in 1922,\(^{19}\) that he has gained “the impression that this consideration is very probably correct from a review article of Compton”, and discusses the case that “the primary X-ray of frequency \(\nu_0\) transfers an energy \(h\nu_0\) to a free electron in a single process”, on the basis of the theory of light quantum. Debye derives for the frequency \(\nu\) of the scattered X-ray a formula which becomes identical with Eq. (8) when expressed in wavelength \(\lambda\). About the angular distribution of the intensity of the scattered X-ray, he assumes that it is given by the product of that in the Thomson scattering with \(\nu/\nu_0\), because this product gives the result “standing on the correspondence principle” for long waves of \(\nu/\nu_0 \simeq 1\), and weaker intensities for short waves of \(\nu/\nu_0 < 1\).

§3. The theory of light quantum and the Bragg formula for X-ray reflection

Nearly at the same time as the application of the theory of light quantum to the scattering of X-rays by electron is made by Compton and by Debye, the Bragg formula for X-ray reflection is derived from the standpoint of the theory of light quantum by Duane, in his paper received on Mar. 2 and published in May in 1923, “The Transfer in Quanta of Radiation Momentum to Matter”.\(^{20}\) As is well known, the Bragg formula is based on the electromagnetic wave theory of X-rays and derived by taking into consideration the effect of interference among X-rays scattered by different lattice points of a crystal.

Previously, in collaboration with Clark, Duane\(^{21}\) carries out experiments in 1922 on the reflection of X-rays from crystals, by making use of a X-ray tube with a tungsten target to which high voltages are applied. The X-ray beam used contains rays of a great variety of wavelengths, but in a specified direction \(\theta\) such rays as satisfy the Bragg formula

\[
n\lambda = 2d \sin \theta
\]

are reflected (\(n\) and \(d\) being an integer and the lattice spacing, respectively). To select a single ray of a definite wavelength and thereby to determine \(d\), they make use of the fact that in the course of reducing the applied voltage, at a

\(^{19}\)A. H. Compton, the same as Ref. 12).

\(^{20}\)W. Duane, Nat. Acad. Sci. 9 (1923), 158.

\(^{21}\)G. L. Clark and W. Duane, ibid. 8 (1922), 90; 9 (1923), 126, 131.
critical value of $V$ the intensity of the reflected X-ray of wavelength given by $Ve = h(c/\lambda)$ vanishes.

They found, for the case of the crystal of potassium iodide (KI), besides "the normally reflected rays", namely those obeying the Bragg formula, "the abnormally reflected rays" from the (100) and (110) planes of the crystal. From the critical values of $V$, the abnormally reflected rays seem to be due to the characteristic K-line of iodine, which is excited by the primary beam. The abnormally reflected rays are found in specified directions different from but near those of the normally reflected rays. About these results Clark-Duane suppose that "if we adopt the interference of scattered rays as the explanation of X-ray reflection, we must assume that, on the average at least, definite time relations must exist between the passage of the primary X-ray across an atom and the emission of a secondary characteristic ray", for the reason that with such time duration reflections in directions different from that given by the Bragg formula could take place.

Contrary to this supposition Duane thinks, in his 1923 paper cited above, that "the reflection by a crystal of X-radiation characteristic of the atoms in a crystal itself does not appear to be explainable in a simple manner by the theory of interference of waves". He thus turns to treating not only the abnormal reflection but also the normal reflection of X-rays on the basis of the theory of light quantum. Duane says that "the fundamental hypothesis of the theory now presented is that the momentum of radiation is transferred to and from matter in quanta, and, further, that the laws of the conservation of energy and of momentum apply to these transfers".

For the sake of simplicity, the axes of the crystal are assumed to lie at right angles to each other along the directions of the axes of the coordinate system. The direction of the primary X-ray is supposed to be parallel to the $(x, y)$ plane, and make an angle $\theta$ with the $x$-axis. The lattice spacings in the $x$- and $y$-directions are denoted by $a$ and $b$, respectively. The primary X-ray is considered to be of energy $h\nu$ and momentum $h\nu/c$. "At some point in its path the X-ray may be deflected" so as to travel in a direction, making an angle $\theta'$ with the $x$-axis. At this point of deflection, the X-ray transfers some energy and momentum to the crystal. Since the mass of the crystal is so large that the velocity acquired by it must be very small, and consequently "the energy transferred from the radiation to the crystal may be neglected".

The momentum transferred to the crystal is "to be delivered in quanta" and is "therefore proportional to the constant $h$". From dimensional reasoning,
the proportionality coefficient must have the dimensions of the inverse of a length. Since for crystalline phenomena the lattice spacings are quantities of importance, the momentum transferred in the \(x\)-direction, \(M_x\), may be put in the form \(M_x = \tau h/a\) with an integer \(\tau\). Similarly, \(M_y = \tau' h/b\) with another integer \(\tau'\).

"Another way of looking at these equations is to regard them as peculiar applications to the transfers of momenta of the quantum equation \(\int pdq = nh\)."\(^{22}\) Saying so, Duane attempts to apply the quantum condition to the momenta which the crystal could take. The condition applied to \(M_x\) and \(M_y\) gives the equations

\[
(h\nu/c)(\cos \theta - \cos \theta') = \frac{\tau h}{a}, \quad (h\nu/c)(\sin \theta - \sin \theta') = \frac{\tau' h}{b}
\]

In the case of \(\tau = 0\) and \(\tau' \neq 0\), it turns out therefore that \(\theta = -\theta'\). With the use of the relation, \(\nu/c = 1/\lambda\), there results the equation

\[
\tau' \lambda = 2b \sin \theta.
\]

This is the Bragg formula for reflection from the principal plane of the crystal normal to the \(y\)-axis.

In this way, Duane arrives at the Bragg formula, which has been based on the wave theory, from the standpoint of the theory of light quantum, through the mediation of the quantum condition for the change in momentum of the crystal. We have seen, previously in Sec. 4, §1, Chap. 3 of the present volume, Rubinowicz’s work in 1918. In this work Rubinowicz showed that the angular momentum \(\Delta J\) and energy \(\Delta E\) of the radiation emitted by an electron during one period of its motion in a circular orbit satisfy the relation \(|\Delta J| = \Delta W/2\pi \nu\), which, together with the relation \(|\Delta J| = h/2\pi\) given by the quantum condition for the transition of the electron, gives the equation \(\Delta W = h\nu\) representing the quantum nature of light. We have seen also in the same place that to Sommerfeld in 1922 “this mediating standpoint seems satisfactory”, but to him in 1924 “Compton’s research seems to exclude such a mediation”. However, as is described just above, at nearly the same time as the discovery of the Compton effect, a “mediating standpoint” of deriving reversely the wave nature of light form the quantum nature of light is presented.

In spite of such a mediating standpoint, Duane mentions that the reasoning by which the Bragg formula is derived above “cannot be considered a logical

\(^{22}\)The italic is due to the present authors.
demonstration”, but “we may regard the reasoning as a means of suggesting equations to be tested by applying them to experimentally determined fact”, because of the conflict between the wave- and particle-natures of light. With this consideration, Duane expresses the energy and momentum of X-ray by \( \varepsilon \) and \( \varepsilon/c \), instead of \( h\nu \) and \( hc/c \), respectively, and rewrites the equations under consideration. He puts then \( \varepsilon = hc/\lambda \) by supposing that “\( \lambda \) denotes some parameter connected with the radiation quantum \( \varepsilon \), but need not be regarded in our present theory as the distance between waves”. His supposition is not reasonable, however, because the relation \( \varepsilon = hc/\lambda \) does not hold unless it concerns a wave of a definite wavelength \( \lambda \). It should be noted that the conflict between the wave- and particle-natures of light consists in the fact that light of frequency \( \nu = c/\lambda \) consists of the light quanta of energy \( h\nu \).

Concerning the reflection of characteristic rays of the atoms in the crystal, Duane makes the assumption that the incident quantum of radiation delivers part of its momentum without changing its direction to an atom and an electron to excite the characteristic rays, and the rest of the momentum is transferred to the crystal as in the former case. On a further assumption that the characteristic rays almost consist of the K-lines of energy \( \varepsilon_1 \), he gets approximately

\[
(\varepsilon_1/c)(l_k - l_k') = \frac{\tau_k h}{d_k}; \quad (k = 1, 2, 3).
\]

In these equations, the \( l_k \)'s denote the direction cosines of the incident X-ray, the \( l_k' \)'s those of the characteristic X-ray, the \( d_k \)'s the lattice spacings and the \( \tau_k \)'s integers. According to these equations, the characteristic X-ray is emitted in directions near those of the reflected ray obeying the Bragg formula.

Now, Compton discusses, in his paper received on Aug. 5 and published in Nov., 1923 “The Quantum Integral and Diffraction by a Crystal”, the problem of the passage of radiation through a crystal,\(^{23}\) “according to the method suggested by Duane, though in somewhat greater detail”. The argument used here “is precisely similar to” that in his paper on the Compton effect, which we have mentioned in the foregoing subsection.\(^{24}\) This is because the incident X-ray transfers its momentum to the crystal, instead of the electron in the case of the Compton effect.

\(^{23}\)A. H. Compton, Nat. Acad. Sci. 9 (1923), 359.
\(^{24}\)A. H. Compton, Ref. 6) of the present chapter.
In doing so, Compton puts the quantum condition for the crystal in the form

$$\bar{p} = \int p dq / \int dq = \frac{nh}{q_1} + \gamma,$$

(15)

where $\bar{p}$ is "the displacement average of the momentum", $q_1 = \int dq$ is "the displacement necessary to bring the system back to its original condition", and $\gamma$ is "the constant corresponding for example to zero energy of an oscillator according to Planck's second radiation formula".\(^{25}\) If $a_x$ is the lattice spacing in the $x$-direction, one has $q_1 = a_x$ and therefore $\bar{p}_x = n_x h / a_x + \gamma_x$. "Since $p_x = \bar{p}_x$ for uniform motion", the change in momentum of the crystal along the $x$-direction is given by $\delta p_x = (h / a_x) \delta n_x$.

Applying this formulation also to "the case of a beam of infinite plane waves of wavelength $\lambda$", Compton considers that "it is clear that after the beam has propagated itself through a complete wavelength it is again in its original condition", and puts $q_1 = \lambda$. The momentum of the beam in the direction of propagation becomes therefore $p = nh / \lambda + \gamma$. This corresponds, "according to the relativity theory", to an energy $\epsilon = pc = nhc / \lambda + \gamma c = nh\nu + \gamma c$. If $n = 1$ and $\gamma = 0$, one has $\epsilon = h\nu$, "which is in accord with the results of photoelectric experiments", and therefore "the momentum of the light is $p = h / \lambda$".

Compton's formulation reminds us of Rubinowicz's formulation in 1917, which we have described in Sec. 4, §1, Chap. 3 of the present volume. Rubinowicz expressed the electromagnetic field in a cavity as the superposition of the eigen oscillations in the cavity, and derived, by applying the quantum condition to each eigen oscillation, the equation $E = \sum n_i h\nu_i$ for the energy of the electromagnetic field in the cavity. As we have mentioned there, this acts as the germ of the quantization of wave fields. Compton's formulation, on the other hand, acts as the germ of the quantization of wave fields with periodic boundary condition. Compton states as follows — "The present quantum conception of diffraction is far from being in conflict with the wave theory. In fact we were able to quantize the incident radiation only in view of the fact that it repeats itself at regular space intervals. Thus even from the quantum viewpoint electromagnetic radiation is seen to consist of waves" —.

Epstein and Ehrenfest say, in their 1924 paper "The Quantum Theory of the Fraunhofer Diffraction",\(^{26}\) that "Compton has recently pointed out

\(^{25}\)This indicates Planck's attempt to found his radiation formula, in which he used as the energy of the oscillator the average $(n + 1/2)h\nu$ of $nh\nu$ and $(n + 1)h\nu$ [M. Planck, Verhand, Deuts. Phys. Gesell. 13 (1911), 138].

that Duane's hypothesis can be justified by the application of the general rules of the theory of quanta to the translatory motions of a crystal lattice". However, because Duane and Compton deal with the case of infinite lattices, Epstein-Ehrenfest study the problems with respect to finite gratings. For this purpose, they consider that in the crystal "the material points with which light quanta may collide are distributed with a certain density $\rho$", which is called the "electronic density" for short. In the case of infinite lattices, if the grating is moving at a constant velocity, the electronic density will return to its original value after the grating has moved through the lattice spacing. For the case of finite lattices, they first carry out the Fourier integral decomposition of $\rho$, and apply Duane-Compton's method to each component in the decomposition.

2. Light Quantum Theory and de Broglie's Wave Theory

§1. Quantum condition and Brillouin's theory of inner atomic elastic wave

The orbits of atomic electrons are quantized by the quantum condition, and are prescribed by quantum numbers. In order to understand the appearance of quantum numbers, Brillouin attempts to introduce a kind of wave, in his paper in 1919 "Mechanical actions hereditable discontinuously by propagation: Trial of the quantum dynamical theory of the atom".\(^{27}\) This is because "the discontinuity comes rather inversely from the motion of continuous matter".\(^{28}\)

Brillouin considers "a particle which moves in an elastic medium with a velocity much larger than the speed $\omega$ of elastic waves". He supposes that "either by proper vibrations or because of the result of displacement in the medium", the particle at every time emits waves emanating from its instant position as the center. If the trajectory is periodic or quasi-periodic, the particle will rejoin a finite number of times the waves which it has emitted in the course of its previous motion. If at the time $t$ and at the point $M(x, y, z)$ the particle rejoins the wave emitted by it at the time $t - \tau_k$ and at the point $M_k(\xi_k, \eta_k, \zeta_k)$, one has

$$\tau_k = \omega \int_{M_k}^{M} \frac{ds}{u} = \omega \tau_k, \quad (16)$$

\(^{27}\)M. Brillouin, Compt. Rend. 168 (1919), 1318.

where \( r_k = \{(x - \xi_k)^2 + (y - \eta_k)^2 + (z - \zeta_k)^2\}^{1/2} \), \( ds \) is the infinitesimal element of the orbit, and \( u \) is the velocity of the particle along the trajectory.

If \( u \) is much larger than \( \omega \), a finite number of the previous points \( M_k \)'s correspond to each point \( M \). “This constitutes well a hereditary field (champ à hérédité), following the terminology of Volterra, but discontinuous.” The particle moves therefore in “the field of action of \( n \) virtual points, which it trails behind it on its trajectory”. Brillouin says that “it appears very evident that the permanent motions are those of constant number \( n \)”, and considers that “the energy of the particle in the field of its \( n \) previous positions experiences a finite variation when the number \( n \) changes by a unit”.

Brillouin presents thus the following as “a dynamical hypothesis given with the necessary qualities to represent the essential properties of the atom of Bohr” — “Besides the velocity of light, the universal medium (aether) possesses a much smaller speed of propagation [of the order of several tens of kilometers per second (?)]. The phenomena of quanta appear when the electron moves with a velocity larger than that speed, along a quasi-periodic orbit of the sort that the electron is at every moment in the field of a definite number of its previous positions” —.

He supposes “in particular” as the second hypothesis that “the energy of a mobile point in the field of one of its previous points is \( \Phi = B\beta^2 / r_k \)”. It results therefore from Eq. (16) that

\[
\Phi = \int_{M_k} ds = \frac{B^2}{\omega} .
\]  

(17)

The left-hand side of this equation has the dimension of action, and this equation shows that this action has a constant value. Brillouin says that “this is the Planck constant \( h \)”. By this consideration, Brillouin attempts to connect the integer \( n \) with the Planck constant, that is, to understand the quantum condition.

As will be shown in the following subsection, Brillouin’s idea of a universal wave having a speed of propagation different from the velocity of light, is taken over by de Broglie in his attempt to interpret the Planck formula of radiation from the view point of the theory of light quantum, bringing forth de Broglie’s wave theory.
§2. From the theory of light quantum to de Broglie’s wave theory

de Broglie, in his 1922 paper “On the Interference and the Theory of Light Quantum”,\textsuperscript{29} attempts to interpret the Planck Formula of radiation by considering “clusters of atoms of light” or “molecules of light”. The fluctuation in the energy density of the cavity radiation $\bar{\varepsilon}^2$ is given, according to Einstein, by Eq. (34) in Chap. 1 of the present volume. Putting $E = V \rho \lambda d\lambda$ in terms of the volume of the cavity $V$, one has

$$\bar{\varepsilon}^2 = h\nu E + \frac{c^3}{8\pi\nu^2 d\nu} \frac{E^2}{V}. \quad (18)$$

$E$ obeys the Planck formula of radiation

$$E = \frac{8\pi hV}{c^3} \frac{\nu^3}{e^{\hbar\nu/kT} - 1} d\nu. \quad (19)$$

de Broglie expands Eq. (19) as a power series in $e^{-\hbar\nu/kT}$ to put

$$E = \sum_n \left( \frac{8\pi hV}{c^3} \right) \nu^3 e^{-\hbar\nu/kT} d\nu = E_1 + E_2 + \cdots + E_n + \cdots, \quad (20)$$

and expresses $\bar{\varepsilon}^2$ as

$$\bar{\varepsilon}^2 = h\nu E_1 + 2h\nu E_2 + \cdots = \sum_n n\hbar\nu E_n. \quad (21)$$

He interprets that “$E_1$ corresponds to the energy divided in quanta of $\hbar\nu$, $E_2$ the energy divided in quanta of $2\hbar\nu$ (molecules of light of 2 atoms), and so on”. Now, the second term in the right-hand side of Eq. (18) represents the fluctuation given by the wave theory. de Broglie thus thinks that “from the viewpoint of light quantum the phenomena of interference appear to be connected with the existence of clusters of atoms”. He says that “it is natural to suppose that if the theory of light quantum succeeds one day in interpreting the interference, it must make such clusters of quanta intervene”. In the introductory part of the short paper under discussion, de Broglie mentions also that “when this synthesis is done, the Maxwell equations will appear without doubt as a continuous approximation to the discontinuous structure of radiated energy, just as the continuous equations in hydrodynamics represent, in a satisfactory way, motions of fluid at our scale”.

\textsuperscript{29}L. de Broglie, Compt. Rend. 175 (1922), 811.
In another 1922 paper "Black Body Radiation and Light Quanta", de Broglie treats the black body radiation by considering it "like a gas formed of atoms of light of energy $W = h\nu$". In this paper, he "neglects molecules of light of 2, 3, \ldots, n$ atoms", so that he "must arrive at the law of radiation of Wien" according to the viewpoint in his paper discussed in the preceding subsection. A body of proper mass $m_0$ moving with a velocity $u = \beta c$ has a kinetic energy $W = m_0 c^2 (1/\sqrt{1 - \beta^2} - 1)$ and a momentum $G = m_0 u / \sqrt{1 - \beta^2}$. "But for the atom of light, $m_0$ must be infinitely small and $\beta$ infinitely near to 1 in such a way that $m_0 / \sqrt{1 - \beta^2}$ would have a finite value $m$." One has then $W = mc^2$ and $G = mc = W/c$. In consideration of $W = h\nu$, therefore, "the mass of the atom is supposed to be equal to $h\nu/c^2$", and "the momentum is $h\nu/c = W/c$".

The number $dn_W$ of atoms of light having momenta of magnitude between $G$ and $G + dG$ is proportional to $e^{-W/kT} 4\pi G^2 dG$. Inserting the relation $G = W/c$, one has

$$dn_W = \text{const.} \times e^{-W/kT} W^2 dW.$$  (22)

Using the usual methods of statistical mechanics, de Broglie calculates the constant factor in this equation, and derives Wien's distribution formula of radiation by putting $W = h\nu$.

"One can also succeed to conceive the gas of atoms of light in the following way." Saying so, de Broglie considers a perfect gas of atoms having energy $W = m_0 c^2 (1/\sqrt{1 - \beta^2} - 1)$ and momentum $G = m_0 u / \sqrt{1 - \beta^2}$. Expressing $4\pi G^2 dG$ in terms of $W$ with the use of these equations for $W$ and $G$, he gets, instead of Eq. (22),

$$dn_W = \text{const.} \times e^{-W/kT} m_0 c^2 \sqrt{\alpha(\alpha + 2)(\alpha + 1)} dW,$$  (22')

where $\alpha = W/m_0 c^2$. If $\alpha$ is small compared with 1, this equation gives the Maxwell distribution law. On the other hand, if $\alpha$ is large compared with 1, this equation gives Eq. (22) again. One arrives at Wien’s distribution law of radiation, “supposing the mass $m_0$ is very small”. In this case, the atom of light moves at “a velocity varying with its energy (frequency), but very extremely close to $c$”.

In this way, de Broglie proposes a rest mass for the light quantum, even if it is assumed to be very small ($< 10^{-50}$ gr. according to his next paper), in his attempt to interpret the law of black body radiation basing thoroughly

---

30) L. de Broglie, Journ. de Phys. 3 (1922), 422.
on the theory of light quantum. However, since in de Broglie's idea energy is directly connected with the frequency, as we have seen above, the problem of the frequency corresponding to the rest energy brings about a difficulty. In order to solve this difficulty de Broglie goes to introduce, following Brillouin, a kind of wave with a velocity of propagation different from that of light.

de Broglie's first paper in 1923, "Waves and Quanta", starts from an indication of this difficulty. By the way, he adds a footnote to the title of this paper, saying "Of the subject of the present note see M. Brillouin, Comptes Rendes 168 (1919), 1318". Now, de Broglie considers in this paper that a material body of proper mass $m_0$ moving at a velocity $u = \beta c (\beta < 1)$ possesses an "internal energy" equal to $m_0 c^2$. "The quantum principle leads one to attribute this internal energy to a simple periodic phenomenon of frequency $\nu_0$ such that $h \nu_0 = m_0 c^2$." "For the fixed observer, to the total energy of the body will correspond a frequency $\nu = m_0 c^2 / h \sqrt{1 - \beta^2}$.

However, if the fixed observer observes the internal periodic phenomenon, on account of the relativistic lag of time, he will see it oscillating slowly with the frequency $\nu_1 = \nu_0 \sqrt{1 - \beta^2}$, different from $\nu$. Now, let us suppose that at time $t = 0$ the moving body coincides with a wave having the frequency $\nu$ defined above, which propagates with the velocity $c / \beta = c^2 / u$ in the same direction as the moving body. If at $t = 0$ there is agreement between the phase of the wave and that of the internal phenomenon of the moving body, after a time $t = x / u$ when the body has passed a distance $x$, the internal phenomenon is represented by $\sin 2\pi \nu_1 (x / u)$, and the wave by $\sin 2\pi \nu (t - x \beta / c) = \sin 2\pi \nu (1 - \beta^2) x / u$. Because of the relation $\nu_1 = \nu (1 - \beta^2)$, these phases remain still in agreement.

de Broglie then makes the following statement — "The atom of light equivalent in the ratio of its total energy to a radiation of frequency $\nu$ is the seat of an internal periodical phenomenon which, seen by the fixed observer, has at every point of the space the same phase as a wave of frequency $\nu$ propagating in the same direction with a velocity remarkably equal to (though very slightly larger than) the constant called velocity of light" —.

Having obtained such an outlook, de Broglie applies furthermore this idea of "fictitious wave" to the electron and derives the quantum condition. Let us suppose that an electron moving at a constant velocity along a closed orbit is at the point O at time $t = 0$. The wave under consideration started from O propagates along the orbit of the electron, and meets again the electron at the

---

31) L. de Broglie, Compt. Rend. 177 (1923), 507.
point $O'$ at time $t = \tau$ such that $OO' = \beta c \tau$. If the period of the orbit is $T_r$, one has $(c/\beta) \tau = \beta c (\tau + T_r)$, so that $\tau = T_r \beta^2 / (1 - \beta^2)$. On the other hand, the internal phase of the electron changes by $2\pi \nu_1 \tau$.

de Broglie considers now that “it is almost necessary to suppose that the trajectory of the electron is not stable unless the fictitious wave coming to $O'$ meets again the electron in phase with it: the wave of frequency $\nu$ and of velocity $c/\beta$ must be in resonance with the length of the trajectory”. From this supposition he has the relation $2\pi \nu_1 \tau = 2\pi n$ ($n$ being an integer) and therefore the condition

$$m_0 c^2 \left( \frac{\beta^2}{\sqrt{1 - \beta^2}} \right) T_r = nh.$$  \hspace{1cm} (23)

This equation can be rewritten as

$$\int_0^{T_r} (m_0 \sqrt{1 - \beta^2}) (u_x^2 + u_y^2 + u_z^2) dt = nh,$$  \hspace{1cm} (24)

which is in accordance with the quantum condition

$$\int_0^{T_r} (p_x dx + p_y dy + p_z dz) = nh.$$  \hspace{1cm} (25)

If the electron moves in a circular orbit of radius $R$ with an angular velocity $\omega$ and a sufficiently small $\beta$, the above condition becomes the Bohr equation $m_0 \omega R^2 = n(h/2\pi)$. In the considerations described above, $\beta$ is assumed constant. But, de Broglie says that if $\beta$ varies but is small enough the quantum condition can be deduced. For larger values of $\beta$, “the equation becomes more complicated and necessitates a special examination”.

§3. de Broglie’s wave and the diffraction and interference of particle

The wave introduced by de Broglie has a velocity of propagation larger than that of light. But, this velocity “cannot correspond to any transport of energy”. On account of this, he considers the wave as “a fictitious wave associated with the motion of a moving body”. It is important to remark that the quantum condition is derived from the condition of resonance imposed on the fictitious wave associating with the given motion of an electron.
In his second 1923 paper "Quanta of Light, Diffraction and Interference", de Broglie calls this wave "the phase wave". This is because a wave of phase velocity $c/\beta$ and frequency $m_0c^2/h\sqrt{1-\beta^2}$ has group velocity $\beta c$, which is identical with that of the moving body. He tries now to explain the diffraction of light quanta on the basis of the idea of phase wave. Taking the case of a free particle doing uniform motion first, de Broglie postulates that "the particle follows the ray of phase wave, that is to say, the normal to the surface of equal phase". "The moving body follows therefore the rectilinear trajectory determined by the Fermat principle applied to the phase wave, which is combined here with the principle of least action applied to the moving body."

Secondly, for the case of a particle which has to traverse an opening whose size is small compared with the wavelength of the phase wave, de Broglie says that "its trajectory will curve in general like the ray of the diffracted wave". His explanation of the diffraction of light quanta is based, thus, on the consideration of the diffraction of the phase wave from the viewpoint of the general property of wave, and on the nature of the trajectory of a moving body being normal to the surface of equal phase. de Broglie states in fact that "the new principle based on dynamics explains the diffraction of the quanta of light even if their number was small".

debroglie considers the diffraction of each individual light quantum, basing thoroughly on the theory of light quantum. Concerning Young's experiment, he mentions that "in the following space behind the slits, the capacity of atoms of light of photoelectric effect varies at every point according to the state of interference of the phase waves that have traversed the two slits making diffraction", and "there will be bright and dark fringes such as the wave theories anticipate and no matter how weak is the intensity of the incident light". In his paper in 1924 "A Tentative Theory of Light Quanta", de Broglie writes as follows — "We shall then see interference fringes, however small may be the number of diffracted quanta, however feeble may be the incident light intensity" — .

Going further, de Broglie argues in the second 1923 paper that not only the light quantum but also "any moving body could in certain cases make 

---

32) L. de Broglie, Compt. Rend. 177 (1923), 548.
33) The emphasis in italics is due to the original paper.
34) L. de Broglie, Phil Mag. 47 (1924), 446; the one in which his three papers in 1923, inclusive of the third paper to be taken below, are got together in English with some discussions in addition.
diffraction”. “A flux of electrons traversing a very small opening would present phenomena of diffraction.” He emphasizes that “from this side it will perhaps be necessary to seek experimental confirmations of our idea”. This is the first time that diffraction is predicted for the electron.

de Broglie “conceives that the phase wave is to guide the transfer of energy, and this is what can permit the synthesis of the undulation and the quantum”. He says that “the new dynamics of a free material point is to the classical mechanics (including that of Einstein), like the undulatory optics is to the geometrical optics”, and that “by thinking over this, one will see that the proposed synthesis appears to be the logical accomplishment of the comparative development of dynamics and optics since the seventeenth century”.

These points mentioned above are, however, concerned with “free material points”. In his third paper in 1923, “The Quanta, the Kinematic Theory of Gas and the Fermat Principle”, de Broglie treats the case of a body which describes a curved trajectory under a field of force. To assert the accordance between the phases of the wave and the moving body, “one is lead to suppose that the phase wave of a moving body of a given total energy has at every point a frequency and a velocity determined by the value which the velocity of the body at that point would have”. Considering so, he thinks that “an enlarged electromagnetic theory without doubt will give us the mechanism of this complex propagation”. To him, “it seems that we know in advance its principal conclusion that the rays of the phase wave coincide with the possible dynamical trajectories”.

As the grounds for this, de Broglie mentions that the rays of the phase wave are given, as in a dispersive medium, by the Fermat principle

\[
\delta \int \frac{\nu ds}{V} = \delta \int \frac{m_0 \beta c}{h \sqrt{1 - \beta^2}} ds = 0
\]  

(26)

in terms of the frequency \( \nu = m_0 c^2 / h \sqrt{1 - \beta^2} \) and the velocity \( V = c / \beta \) of the phase wave, while the trajectories of a moving body are determined by the principle of least action

\[
\delta \int m_0 c^2 \left( \frac{1}{\sqrt{1 - \beta^2}} - \sqrt{1 - \beta^2} \right) dt = \delta \int \frac{m_0 \beta c}{\sqrt{1 - \beta^2}} ds = 0.
\]  

(27)

\[35\] L. de Broglie, Compt. Rend. 177 (1923), 630.
Of the possible dynamical trajectories, "some possess the particular property of being in resonance with the phase wave, namely, they are Bohr's stable trajectories for which $\int v ds/V$ is an integer".

In this way, de Broglie attempts "the synthesis of the undulation and the quantum" by supposing that the phase wave associated with a moving body leads the body in its orbital motion. But, he takes only the general properties of the phase wave into consideration, not arriving at the presentation of the wave equation. That is, the equation of motion in "the new dynamics" which should correspond to the wave optics in the analogy supposed by him. The Fermat principle is to give an approximation in the sense of the geometrical optics to the propagation of wave, which is effective in the cases that the change in refractive index can be neglected over a distance comparable with the local wavelength. de Broglie says, in his 1924 paper cited above, that "when Fermat' principle is no longer valid for computing the ray path, the principle of least action is no longer valid for computing the body path". The wave equation in "the new dynamics" would be presented by Schrödinger in 1926 for material particles.

Of the light quantum, de Broglie starts his considerations basing entirely on the particle theory, in the sense that a finite though very small mass is assumed for "the atom of light". He then introduces as that which is responsible for the wave nature of light the phase wave, which is associated with an "atom of light" moving at a velocity slightly less than the light velocity, and has a phase velocity slightly greater than the light velocity and a group velocity identical with the velocity of the "atom of light". For this new kind of wave, however, no "enlarged electromagnetic theory" is arrived at.

Needless to say, the "classical theory" in the theory of light should be, in our context, the electromagnetic theory. Because of this fact, we have had the germ of the quantization of field presented by Rubinowicz in 1918, or by Compton in the same year 1923 as the proposal of de Broglie's theory of wave. Here, we see the very reason that de Broglie's theory of wave which is born from the particle theory of light becomes effective in bringing forth the Schrödinger wave equation for material particles with mass.

§4. Theory of light quanta and Bose-Einstein statistics

(i) Bose's statistical theory of gas of light quanta In the third paper in 1923, de Broglie considers that "every wave can transport zero, one, two
or several atoms, of such a kind that follows the canonical distribution”. The
nth term in the right-hand side of Eq. (20) is, thus, supposed to correspond
to “n atoms of light” obeying the canonical distribution to give the Planck
distribution formula. In this third paper de Broglie uses “n atoms of light”
instead of a “molecule of light of n atoms” in his 1922 papers cited above.
In any case, he considers that the black body radiation consists of groups of
various numbers of light quanta.

The derivation of the Planck formula from the standpoint of the theory of
light quantum in a fully statistical way is attempted by Bose and accomplished
by Einstein 1924. At the same time, it is shown by Einstein that fluctuation
in the number of molecules per unit volume in a gas, which is derived with use
of the new statistics, can be explained by the wave nature of material particles
proposed by de Broglie.

Bose, in his paper “Planck's Law and Light Quantum Hypothesis”,36) says
that “all the previous derivations make use of the relation between the density
of radiation and the average energy of an oscillator”. In this relation, the
factor \(8\pi \nu^2 d\nu/c^3\) comes from “the assumption for the number of the degrees
of freedom of the aether”. He considers that “this factor could, however, be
derived only from the classical theory”, and “this is the unsatisfactory point
in all the derivations”.

Bose continues to say that “a remarkable elegant derivation has been given
by Einstein”. From “very simple assumptions for the energy interchange among
molecules and radiation field”, Einstein finds the expression for the density of
radiation \(\rho_\nu\).37) In order to bring this expression in accordance with the Planck
formula, however, “he has to make use of Wien’s displacement law and Bohr’s
correspondence principle”.38) Bose says that “Wien’s law is based on the
classical theory, and the correspondence principle assumes that the quantum
theory agrees with the classical theory in certain limiting cases”.

Bose thus thinks that “in all the cases the derivations seem not to be justi­
ﬁed logically enough”. On the contrary, to him “the light quantum hypothesis

36) Bose, Zeit. f. Phys. 26 (1924), 178 [published on being “translated by A. Einstein” into
German].
37) Cf. Eq. (94) in Chap. 3 of the present volume.
38) Actually, as we have described in Sec. 3, §2, Chap. 3 of the present volume. Einstein
showed the relations \(\alpha_{mn} \propto \nu^3\) and \(\epsilon_m - \epsilon_n \propto \nu\) by making use of Wien’s displacement
law. Namely, he derived Bohr’s frequency relation with the use of Wien’s displacement
law. Bohr’s correspondence principle was proposed later in order to calculate, on the basis
of the classical theory, the transition probability introduced by Einstein in his work under
consideration, as we have also described in the same section.
connected with statistical mechanics seems to be sufficiently independent of the classical theory for the derivation” of Planck’s law. Now, Bose supposes that “the momentum state of a light quantum is characterized by its coordinates $x, y, z$ and the respective momenta $p_x, p_y, p_z$” satisfying the relation $p_x^2 + p_y^2 + p_z^2 = (h\nu/c)^2$. The volume of the phase space corresponding to the frequency interval $d\nu$ is given by $4\pi V(h\nu/c)^2 h d\nu/c$. Dividing this by $h^3$ following Planck and then multiplying it by 2 “to take into account the fact of the polarization”, one gets the number of the statistic mechanical cells corresponding to $d\nu$ as $8\pi V \nu^2 d\nu/c^3$.

The number per unit spatial volume, $8\pi \nu^2 d\nu/c^3$, is the one that appears in the Rayleigh–Jeans distribution formula as “the number of the degrees of freedom of the aether” in Bose’s statement. In Bose’s derivation of this number, the number of cells into which light quanta of momentum $h\nu/c$ are distributed becomes independent of $h$, on account of taking $h^3$ for the volume of an individual cell in the phase space. de Broglie also uses this value of the volume of the cell to calculate Helmholtz’s free energy in the derivation of the Wien distribution formula from the viewpoint of the theory of light quanta in his 1922 paper cited above.

In this way, the number of cells in the phase space allowed to light quanta is independent of the Planck constant $h$ and agrees with the number of the eigen oscillations in the electromagnetic wave theory, even if $h$ is introduced on the basis of the theory of light quantum. At this point, we can see one reason why Planck succeeded phenomenologically to arrive at the Wien radiation formula, and then at the Planck radiation formula, though he based his derivation on the classical theory. This situation is somewhat similar to the case of the Rutherford scattering and of the Zeeman effect. From this viewpoint, too, the idea given by de Broglie and Bose has a great historical significance.

Bose then calculates the number of ways of partitioning the states of light quanta into the cells of the phase space. If there are $N^2$ light quanta of energy $h\nu^s$ ($s = 0, 1, 2, \ldots \infty$), the total energy of the radiation is given by

$$E = \sum_s N^s h\nu^s.$$  \hspace{1cm} (28)

39) The superscript $s$ in $N^s, \nu^s$ and others is an ordinal number to label a cell, not an index number.
For the number of ways of distributing $N^s$ light quanta among the $A^s = (8\pi\nu^s/\epsilon^3) d\nu^s$ cells, Bose uses

$$A^s/p_0^s! p_1^s! \ldots,$$

where $p_0^s, p_1^s, p_2^s, \ldots$ are the numbers of the cells that have respectively $0, 1, 2, \ldots$ light quanta belonging to $N^s$ quanta, so that

$$N^s = 0 \cdot p_0^s + 1 \cdot p_1^s + 2 \cdot p_2^s + \ldots. \quad (30)$$

"The probability of the state defined by all the $p^s_r$'s is clearly $\prod_s (A^s/p_0^s! p_1^s! \ldots)$".

Denoting this by $W$, one has, in consideration of the largeness of $p^s_r$'s,

$$\log W = \sum_s A^s \log A^s - \sum_s \sum_r p^s_r \log p^s_r. \quad (31)$$

On the subsidiary conditions that $E = \sum_s N^s \epsilon^s$ and $N^s = \sum_r r p^s_r$ should be constant, $\log W$ is found to be maximum for $p^s_r$'s such that

$$p^r_s = A^s \left\{ 1 - \exp \left( -\frac{\epsilon^s}{\beta} \right) \right\} \exp \left( -\frac{r \epsilon^s}{\beta} \right), \quad (32)$$

where $\beta$ is a Lagrange multiplier. From the values of $E$ and the entropy $S$ for these values of $p^r_s$'s, $\partial E/\partial S$ is calculated to give the relation $\beta = kT$. It consequently follows that

$$E = \sum_s \frac{8\pi \nu^s}{\epsilon^3} \frac{1}{\exp(h\epsilon^s/kT) - 1} d\nu^s, \quad (33)$$

Bose concludes that "this equation is equivalent to the Planck formula".

(ii) **Einstein's quantum statistical theory of gas** Einstein immediately uses Bose's method in a 1924 paper "Quantum Theory of Monoatomic Ideal Gas".\(^{40}\) For an ideal gas of monoatomic molecules, the number of cells in the phase space $\Delta s$ corresponding to the energy interval $\Delta \epsilon$ is, instead of Bose's $A^s$,

$$\Delta s = 2\pi \left( \frac{V}{\hbar^3} \right) (2m)^{3/2} \epsilon^{1/2} \Delta \epsilon, \quad (34)$$

where $m$ is the mass of the molecule, in terms of which the energy of a molecule is expressed as $\epsilon = (1/2m)(p_x^2 + p_y^2 + p_z^2)$. Through calculations similar to

those in Bose’s paper, Einstein derives for the total number of molecules $N = \sum_s \sum_r \epsilon^s r p^s_r$ and the total energy $E = \sum_s \sum_r \epsilon^s r p^s_r$ the following expressions:

$$N = \sum_s \frac{\Delta s}{\{\exp(\alpha^s) - 1\}} ,$$  

$$E = c \sum_s \frac{s^{2/3} \Delta s}{\{\exp(\alpha^s) - 1\}} .$$

In these equations, $\epsilon^s$ is the energy of a molecule in a cell with the value of $s$ determined by the relation $\epsilon^s = cs^{2/3}$, which is obtained by integrating Eq. (34), and $\alpha^s$ is defined by

$$\alpha^s = A + \frac{cs^{2/3}}{kT} ; \quad c = \left( \frac{\hbar^2}{2m} \right) \left( \frac{4\pi V}{3} \right)^{-2/3} .$$

The quantity $\lambda = e^{-A}$, defined by the constant $A$ which denotes the second Lagrange multiplier other than $\beta$, takes a small value at high temperatures and approaches unity as the temperature tends to zero. $\lambda$ represents the degree of “degeneracy” of the gas. That is, the energy per molecule $E/N$, which is calculated by expansion in a power series in $\lambda$ of the summands in Eqs. (35) and (36) and substituting the summation $\sum_s$ by the integration $\int ds$, takes a value smaller and smaller than the $(3/2)kT$ due to the classical theory as $\lambda$ approaches unity. At the absolute zero temperature, all the molecules are therefore found in the first cell of zero energy. For this state of the gas “there exists only one single distribution of the molecules”. The entropy of the gas thus vanishes at the absolute zero temperature. Einstein thus arrives at the idea which is nowadays called the Bose–Einstein condensation.

In the second paper “Quantum Theory of Monoatomic Ideal Gas: 2nd Treatise,” received four months after the paper mentioned above, Einstein says that Ehrenfest and his collaborators criticize Bose’s and Einstein’s theories, pointing out that though in both the theories “the light quanta or the molecules are not treated as substances statistically independent of each other”, this circumstance has not been clearly mentioned. Einstein admits that “this point is quite right”.

Although no explicit citation is given by Einstein, Ehrenfest and Kamerlingh-Onnes presented “A simple Derivation of the Formula, which Planck

used as the Ground for his Theory of Radiation" in their 1914 paper with the quotation as title.\footnote{42} Namely, they derived Planck’s number of complexions $(N - 1 + P)!/(N - 1)!$, which Planck took as the number of the ways of distributing $P$ quanta of energy among $N$ oscillators, from the formula of combination with repetition. If $P$ energy quanta are arranged in a line and divided into $N$ groups with $(N - 1)$ partitions, the number of the quanta in each group corresponds to the number of the quanta distributed to each oscillator. The number of ways of arranging the quanta and partitions in a line, $(N - 1 + P)!$, divided by that of arranging the quanta in a line, $P!$ and that of the partitions in a line, $(N - 1)!$, is just equal to Planck’s number of complexions.

Ehrenfest and Kamerlingh-Onnes pointed out that, in Einstein’s paper in 1905 on the theory of light quantum, the number of ways of distributing $P$ light quanta, “identical in kind and free from each other”, among $N$ cells in the space was taken to be $N^p$.\footnote{43} This is equal to that in the case of “$P$ gas molecules identical in kind and independent of each other”, but different from Planck’s number of complexions. In Planck’s case, no supposition was made of “$P$ mutually independent free quanta”. Ehrenfest and Kamerlingh-Onnes came to the following conclusion — “Einstein’s viewpoint leads inevitably to Wien’s law of radiation. … Planck’s formal technique (the distribution of $P$ energy quanta among $N$ oscillators) cannot be interpreted in the meaning of Einstein’s light quanta” —.\footnote{44}

Next, in the second paper under consideration, Einstein makes “comparison between the two ways of thinking”. Replacing $\Delta s$ in the first paper with $z_{\nu}$, Einstein takes as the number of ways of distributing $n_{\nu}$ molecules among $z_{\nu}$ cells

$$w_{\nu} = \frac{(n_{\nu} + z_{\nu} - 1)!}{n_{\nu}!(z_{\nu} - 1)!}.$$  (38)

This represents the way of distribution which is nowadays called the Bose-Einstein statistics. Equation (38) is different from number of the ways of distribution in Bose’s case given in Eq. (29), though no mention is made specif-

\footnote{42}P. Ehrenfest und H. Ksmerlingh-Onnes, Ann. d. Phys. 46 (1915), 1021; the original paper in Dutch was published in 1914.

\footnote{43}Cf. Sec. 3, §7, Chap. 1, Vol. I of this book. It suffices to consider that the volume $V$ of the space full of radiation in that section is $N$ times the volume of a cell.

\footnote{44}It is well known nowadays that when $P$ is small enough compared with $N$, $(N - 1 + P)!/(N - 1)!P! \approx N^p/P!.$
ically in Einstein's paper. In Eq. (38) \( n_\nu \) identical molecules are supposed to be distributed among \( z_\nu \) cells, while in Eq. (29) \( p_\nu^s \) cells of a total of \( A^s (= \sum_r p_r^s) \) cells are supposed to be identical. The entropy resulting from the logarithm of \( W = \prod_\nu w_\nu \) is given by

\[
S = k \sum_\nu \left( (n_\nu + z_\nu) \log(n_\nu + z_\nu) - n_\nu \log n_\nu - z_\nu \log z_\nu \right).
\] (39)

On the other hand, if the molecules are statistically independent of each other, the number of ways of distributing \( n_\nu \) molecules among \( z_\nu \) cells is equal to \( z_\nu^{n_\nu} \). Since the order of molecules in every cell does not matter, \( w_\nu \) in this case is given by

\[
w_\nu = \left( \frac{n!}{\prod_\nu n_\nu!} \right) z_\nu^{n_\nu}; \quad n = \sum_\nu n_\nu, \quad (38')
\]

instead of Eq. (38). This represents a way of distribution that is nowadays called the Maxwell-Boltzmann statistics. This gives

\[
S = k \sum_\nu n_\nu (\log z_\nu - \log n_\nu). \quad (39')
\]

By making \( S \) given by Eq. (38) or (38') maximum on the subsidiary conditions \( E = \sum n_\nu E_\nu = \text{const.} \) and \( n = \sum n_\nu = \text{const.} \), Einstein gets the expression for \( n_\nu \) for the respective cases,

\[
n_\nu = z_\nu \left\{ \exp \left( \alpha + \frac{E}{kT} \right) - 1 \right\}, \quad (40)
\]

and

\[
n_\nu = z_\nu \exp \left( -\alpha - \frac{E}{kT} \right). \quad (40')
\]

(iii) The B-E statistics and wave-nature of material particles A comparison of Eq. (40) and Eq. (40') "suggests that we are on the assumption that the molecules are under mutual influence in a manner nonunderstandable at present". Saying so, Einstein stresses that "this influence is just the condition for the various states which are called complexions to have equal statistical

\[\text{[45] Though Einstein made here no mention of the Planck formula, for the gas of light quanta one has } \alpha = 0 \text{ because the subsidiary condition } n = \text{const. is not imposed on it, as is seen in text books on the statistical mechanics. By the way, Bose imposed the condition } N^s = \text{const. (} n_\nu \text{ = const. in Einstein's corresponding notation), as we have seen above.} \]
probability”. If one calculates the mean-square of the fluctuation $\Delta_{n}$ of $n_{e}$ by making use of $S$, one gets from Eq. (39)

$$\left(\frac{\Delta_{n}}{n_{e}}\right)^{2} = \frac{1}{n_{e}} + \frac{1}{z_{n}}. \quad (41)$$

On the other hand, from Eq. (39') one gets only the first term in the right-hand side of this equation. The second term represents, in the case of cavity radiation, the value due to the classical theory, that is, the value resulting from the wave-nature of radiation.

Accordingly, Einstein takes notice of de Broglie’s theory of wave. de Broglie supposes an internal harmonic oscillation of frequency $\nu_{0} = mc^{2}/h$ for a material particle of mass $m$ at rest. Einstein derives de Broglie’s phase wave by considering that, in the coordinate system in which the particle moves at velocity $u$, the harmonic oscillation is seen as the wave motion

$$\sin \left\{ 2\pi \nu_{0} \frac{(t - ux/c^{2})}{\sqrt{1 - u^{2}/c^{2}}} \right\}, \quad (42)$$

because of the transformation of time. Einstein states that the second term $1/z_{n}$ in the expression for $\left(\Delta_{n}/n_{e}\right)^{2}$ represents “the fluctuation of this wave field”, because according to de Broglie’s theory “a scalar wave field is made to correspond to a gas”. This statement should be said to be the first confirmation of the wave-nature of material particles, though this confirmation is made through a theoretical analysis. The cause by which “the molecules are under mutual influence in a manner nonunderstandable at present” is thus grasped by means of the wave-nature of molecules.

When $u/c$ is small, the wavelength $\lambda$ of de Broglie’s phase wave becomes $\lambda = h/mu$, which usually takes a value very small compared with the diameter $\sigma$ of the molecule concerned. For hydrogen or helium gas, however, $\lambda$ becomes of the same order of magnitude as $\sigma$ at low temperatures, so that phenomena of diffraction may take place in intermolecular collisions. Einstein points out that as a result “scatterings of molecules nonunderstandable mechanically” participate in the molecular process in the gas, with a reduction in the mean free path of molecules, and therefore “a somewhat sudden decrease in the viscosity occurs as the temperature becomes low”. As a rough estimation of such temperatures, he gives 56 K for $\text{H}_2$ and 40 K for He.
In addition, Einstein argues that, though at low temperatures a gas of molecules could not be treated as an ideal gas because of the increase in the apparent molecular size due to the effect of diffraction, free electrons in a metal may be treated as an ideal gas. With this consideration he says that the difficulty that at normal temperatures free electrons in a metal make almost no contribution to its specific heat "is solved on the basis of the above mentioned theory of gas". That is, he supposes that almost all the free electrons in a metal are in the first cell of zero energy. The real solution to this difficulty is, however, achieved later by the idea of the degenerate gas of electrons based on the Fermi-Dirac statistics. The presentation by Pauli of the exclusion principle for electrons, which we have described in the preceding chapter, is made at nearly the same time as Einstein's second paper under consideration.

Now, taking notice of Einstein's paper, Elsasser at I. Franck of Göttingen University points out that some of the experimental results already obtained are understandable by the notion of the wave-nature of electron. In his paper in 1925, "Remarks on Quantum Mechanics of Free Electron", Elsasser quotes the finding by Ramsauer and others of an anomalous mode of variation with velocity of the mean free path of slow electrons in a gas. Elsasser says that the curve of this variation shows an entirely conspicuous agreement with the curve that is given in the classical theory for the inflexion of light by small colloidal spheres. He notes that it suffices qualitatively to apply the wave length of light to the de Broglie wave length of electron, and the radius of the colloidal particle to that of the atom.

Elsasser makes also an analyses of Davisson and Kunsman's experiment in which a number of peaks are found in the angular distribution of the intensity of electrons reflected by a plate of platinum. The peaks are observed to move in one direction with the velocity of electrons. Elsasser treats the peaks like the

---

46) E. Fermi, taking notice of Stoner's and Pauli's works, which we have seen in the preceding chapter, modified the method of calculation of \( w_n \) and arrived at a new kind of statistics [Zeit. f. Phys. 36 (1926), 902]. P. A. M. Dirac showed, independently of Fermi, in his paper five months later than Fermi's, that the wave function for a system consisting of identical particles should be either symmetric or anti-symmetric for the interchange of every pair of particles, and that the former possibility gives the Bose-Einstein statistics and the later, the Pauli exclusion principle. Compare also Sec. 1, §4, Chap. 3, Vol. III.

47) W. Elsasser, Die Naturwiss. 13 (1925), 711.

48) C. Ramsauer, Ann. d. Phys. 72 (1923), 345. See also J. S. Townsent and V. A. Bailey, Phil. Mag. 43 (1922), 593. The phenomenon found in these works is nowadays called the Ramsauer-Townsent effect.

diffraction pattern of light with the use of the de Broglie wavelength, getting the angles of the peaks in a rough agreement with the experimental ones. Elsasser's work should be said to have given for the first time experimental evidence for the wave-nature of electrons.

(iv) Substantialistic background of the B-E statistics The number of the ways of distribution \( w_\nu \) and the entropy \( S \) given in Eqs. (38) and (39), respectively, which Einstein uses in the second paper, are equivalent as mathematical expressions to those used by Debye in 1910.\(^{50}\) In making \( S \) maximum, Debye imposed only the subsidiary condition that the total energy \( E \) should be constant. In Einstein's case, if the subsidiary condition for the constancy of \( \sum n_\nu \) were not imposed, there would result the Planck distribution formula, too. Einstein's expression of \( w_\nu \) is identical in its mathematical form with Planck's expression of the number of complexions, on which Debye's expression of \( w_\nu \) bases. As we have mentioned in Sec. 1, §3, Chap. 1 of the present volume, Einstein had a critical view of the application of the number of complexions by Planck. Einstein makes use of it on the premise that "the molecules are not treated as substances statistically independent of each other". In Planck's case it was introduced as a technique to interpret his distribution formula.

In the course from Debye in 1910 to Einstein in 1924, there were various developments in the substantialistic knowledge of radiation and atoms. Bose took as the starting point of his study Einstein's theory of radiation in 1916. This theory of Einstein, in which the Planck distribution was derived from a consideration of the equilibrium of energy and momentum among atoms and light quanta in terms of the transition probability newly introduced, was based on the Bohr model of atom, the quantum condition due to Sommerfeld, and the analyses of spectral lines which were made on the basis of these two. The concept of transition probability became possible to be applied concretely to the analyses of spectral lines by virtue of Bohr's correspondence principle.

To Debye in 1910, it seemed that his derivation of the Planck distribution formula by making use of a mathematical expression, which is now seen to be equivalent to that used in the Bose-Einstein statistics, "does not explain the question whether the existence of the elementary quantum should be seen as the characteristic of the aether or as the characteristic of matter". To Bose

\(^{50}\) Cf. Sec. 1, §3, Chap. 1 of the present volume. Note that in Debye's case 1 is neglected in comparison with \( z_\nu \) in the expression of \( w_\nu \), and the constant term \(-k \sum z_\nu \log z_\nu \) is omitted from that of \( S \).
in 1924, as well as to de Broglie in 1923, however, the theory of light quanta seemed to be fully acceptable. Debye also treated, in 1923 at almost the same time as Compton, the scattering of X-ray on a free electron on the basis of the theory of light quantum. Furthermore, Einstein and Debye applied Planck's formula for the average energy of oscillators to the vibrations of molecules in a solid, to elucidate the specific heat of solid bodies at low temperatures. "It never holds that 'equation is everything'".  

3. Bohr-Kramers-Slater's Theory and the Negation by Compton-Simon's Experiment

§1. Slater's idea of virtual radiation

In 1923, Compton analyzed the scattering of X-ray on a free electron, Duane derived the Bragg formula for the reflection of X-rays by a crystal, and de Broglie arrived at the idea of the phase wave, each being on the basis of the theory of light quantum. In the very same year, Bohr expressed some doubts about the theory of light quanta in his review article "On the Application of the Quantum Theory to Atomic Structure".  

In "Chap. III On the formal Nature of Quantum Theory; §1. The Light Quantum Hypothesis" of this review article, Bohr says that "in the so-called 'hypothesis of light quanta' of Einstein", "one, in the first place, lays the weight on the strict retention of such general laws as the conservation of momentum and of energy even in every individual process". "Quite apart from the great importance which, as is well known, this viewpoint has had in placing certain classes of phenomena, such as the photoelectric effect, in a clear light in their relation to the quantum theory, the hypothesis under consideration can, however, by no means be regarded as a satisfactory solution. As is well known, this hypothesis actually brings in insuperable difficulties, when the explanation of the phenomena of interference it dealt with." Saying so, Bohr cites Lorentz's paper in 1910.  

Bohr mentioned further that "we can even maintain that the picture, on which the hypothesis of light quanta bases, excludes in principle the possibility

---

52) N. Bohr, Zeit. f. Phys. 13 (1923), 117 [the same as Ref. 35) in Sec. 1, §4, Chap. 4 of the present volume.
53) Cf. Sec. 3, §3, Chap. 1 of the present volume.
of a meaningful definition of the concept of the frequency $\nu$, which plays just a main role in this theory". That is, he indicates the fact that the frequency $\nu$ in the theory of light quanta comes from the wave theory of light. About these words of Bohr, however, we should note that this fact represents the conflict between the particle- and wave-natures. The theory of light quanta is not a mere return to Newton's particle theory of light.

He continues to say the following — "The appropriate way in which the hypothesis reproduces certain aspects of the phenomena is rather suitable to support the view held from various sides that, in opposition to the description of natural phenomena in classical mechanics where the statistical result of a great number of individual processes is always dealt with, a complete space-time description of atomic occurrences is not practicable in a noncontradictory way with the help of the concepts borrowed from classical electrodynamics, which up to now are our only means of formulating the principles on which the actual application of the quantum theory bases". He states further, in "§4. Law of Energy-Momentum in the Quantum Theory" of the same chapter as above, that "for that reason, we must be prepared for the fact that conclusions from these laws will possess no unlimited validity". Bohr doubts, in this way, the conservation of energy and momentum in atomic elementary processes as well as the theory of light quanta.

About "the field of radiation in an enclosure with reflecting walls", Bohr mentions in "§2. The Coupling Principle" of the same chapter, that because "according to the classical electrodynamics such a field possesses a formal analogy with the motion of a multiple periodic system consisting of material particles", "the energy which goes into each eigen-oscillation (of the field of radiation) is equal, as is immediately clear, to a whole multiple of $h\nu$, where $\nu$ means the frequency of the eigen-oscillation in question". He says that "this idea forms the basis" of Debye's derivation in 1910 of the Planck formula. This description of Bohr is not correct, however, as a historical note. As we have seen in Sec. 4, §1, Chap. 3 of the present volume, it was Rubinowicz who, taking notice of Debye's way of the derivation of the Planck formula, applied in 1917 the quantum condition to each of the eigen-oscillations, of which the field of the cavity radiation is superposed, and showed that the energy of an eigen-oscillation of frequency $\nu$ should be $nh\nu$. At the stage of 1910 this was not "immediately clear".\footnote{Cf. Sec. 1, §3, Chap. 1 of the present volume.}
Any way, Bohr lay stress on the wave-nature of radiation. In "§3. Phenomena of Reflexion and Dispersion" of the same chapter, his discussion goes on — "Our whole knowledge of the properties of radiation, which to some extent plays a decisive role in the problem of the atomic structure, rests of course only on these phenomena, in a closer consideration of which the formal nature of the quantum theory in particular comes in sight clearly". "Thus it seems necessary, in order to carry out the calculations for phenomena of reflection and dispersion, to assume that an atom reacts to the field of radiation in a similar way as a system of charged particles in the classical theory, namely that the atom forms the starting point of a secondary wave train, which stands in a coherent phase relation with the original field of radiation."

"The phenomena of dispersion in gases show that the course of dispersion can be described with close approximation on the basis of a comparison with a system of harmonic oscillators in accordance with the classical theory of electron, if one equates the frequencies of these oscillators to the frequencies of the lines of the observed absorption spectrum of the gas concerned." "On the other hand, according to the postulates of the quantum theory, the frequencies of these absorption lines \cdots{} are determined in accordance with the frequency condition by the difference in energy of the atom in the normal state and in another (excited) state different essentially from it."

"The phenomena of dispersion must therefore be so grasped that the reaction of the atom to radiation is closely related to some unknown mechanism which is responsible for the emission of radiation in the transition." "This mechanism becomes effective when the atom is irradiated in such a way that the total reaction of a number of atoms is the same as that of a number of harmonic oscillators in the classical theory, the frequencies of which are equal to those of the radiation emitted by the atom in the possible processes of passing over, and the relative number of which is determined by the probability of occurrence of such process of passing over in irradiation."

In contrast to Bohr, Slater takes account of the particle-nature of radiation, and presents the idea of "a virtual field of radiation, originating from oscillators having the frequencies of possible quantum transitions", in his paper in 1924 "Radiation and Atoms". He considers here in the following way. "On one hand, interference and dispersion demand a continuity similar to that in the theory of optics. On the other, the exchange of energy and momentum between radiation and matter presents discontinuous features which have even led to

the introduction of the theory of light quanta." "The two aspects must be really consistent, but the discontinuous side is apparently fundamental."

"Nevertheless, on the basis of Bohr's correspondence principle, it seems possible to build up a more adequate picture of optical phenomena than has previously existed, by associating the essentially continuous radiation field with the continuity of existence in stationary states, and the discontinuous changes of energy and momentum with the discontinuous transitions from one state to another."

Slater introduces thus the idea of virtual radiation. Namely, he says that "any atom may be supposed to communicate with other atoms all the time it is in a stationary state, by means of a virtual field of radiation". To him, "the idea of the activity of the stationary states presented here suggested itself in the course of an attempt to combine the elements of the theories of electrodynamics and of light quanta by setting up a field to guide discrete quanta, which might move, for example, along the direction of Poynting's vector".

It may be interesting to note that Slater attempts to grasp the interconnection between the wave- and particle-natures by the idea of "field to guide discrete quanta" which is similar to de Broglie's idea of phase wave, nearly at the same time as de Broglie. The English version of the outline of de Broglie's paper on his wave theory was published in the Oct. 13, 1923 issue of the Nature, and Slater's paper was received on Jan. 28, 1924. In the latter there is no citation of the former.

Now, "when the idea with that interpretation was described to Dr. Kramers" by Slater, Kramers "pointed out that it scarcely suggested the definite coupling between emission and absorption processes which light quanta provide, but rather indicated a much greater independence between transition processes in distant atoms" than Slater has perceived. He says that "the subject has been discussed at length with Prof. Bohr and Dr. Kramers, and a joint paper with them will shortly be published in the Philosophical Magazine". This is Bohr-Kramers-Slater's paper, that we shall discuss in the following subsection to bring out their theory of statistical conservation of energy and momentum.

So, Slater discards the connection of his virtual radiation with light quanta, stating that "the function of the virtual radiation is to provide for statistical conservation of energy and momentum by determining the probabilities for quantum transitions". "The part of the field originating from the given atom itself is supposed to induce a probability that atom loses energy spontaneously,
while radiation from external sources is regarded as inducing additional probability it gains or loses energy, much as Einstein has suggested." In a paper in 1925 "The Nature of Radiation", Slater says that, when his original idea was presented to Bohr and Kramers, "they pointed out that the advantages of this essential feature would be kept, although rejecting the corpuscular theory, by using the field to induce a probability of transition rather than by guiding corpuscular quanta". It may be said that Bohr-Kramers's opinion was a natural one, in consideration of aforementioned Bohr's doubt about the theory of light quantum and the conservation of energy and momentum in an atomic elementary process.

§2. B-K-S theory of statistical conservation of energy-momentum

The paper "The Quantum Theory of Radiation" written by Bohr and Kramers in 1924 in collaboration with Staler has its origin, as we have seen above in §1, in Slater's idea of a virtual field, by means of which an atom communicates with other atoms, and in Kramers' comment on this idea. Bohr, Kramers and Slater say that "the present paper • • • may in various respects be considered as a supplement to the first part of a recent treatise by Bohr". Bohr's paper cited here is the one which we mentioned in the preceding subsection, §1.

They state — "As regards the occurrence of transitions, we abandon any attempt at a causal connection between the transitions in distant atoms, and especially a direct application of the principle of conservation of energy and momentum, so characteristic of the classical theories." "The application of these principles to the interaction between individual atomic systems is, on our view, limited to interactions which take place when the atoms are so close that the forces which would be connected with the radiation field on the classical theory are small compared with the conservative parts of the fields of force originating from the electric charges in the atom. Interactions of this type, which may be termed 'collisions', offer, as is well known, illustrations of the stability of stationary states postulated in the quantum theory."

"Thus we assume that an induced transition in an atom is not directly caused by a transition in a distant atom for which the energy difference between

---

56) J. C. Slater, Nature 116 (1925), 278.
57) N. Bohr, H. A. Kramers and J. C. Slater, Phil. Mag. 47 (1924), 785.
the initial and final stationary state is the same” — namely, they suppose that there is no such process that a light quantum emitted by an atom is absorbed by another atom. They say further — “The degree of independence of the transition processes assumed here would seem the only consistent way of describing the interaction between radiation and atoms by a theory involving probability considerations.”

“This independence reduces not only conservation of energy to a statistical law, but also conservation of momentum.” “If the transition is induced by virtual radiation fields from distant atoms, the direction of the change of the momentum of the atom is the same as that of the wave propagation in this virtual field.” “In the case of a transition by its own virtual transition, we shall naturally assume that the change of momentum is distributed according to probability laws in such a way that changes of momentum due to transitions in other atoms are statistically compensated for any direction in space.” “The cause of the observed statistical conservation of energy and momentum we shall not seek in any departure from the electromagnetic theory of light . . ., but in the peculiarities of the interaction between the virtual field of radiation and illuminated atoms” —.

As we have seen in §1 of the present section, in Bohr’s paper in 1923 the correspondence with a large number of harmonic oscillators in the classical theory is supposed in regard to the reaction of an atom to radiation. In G-K-S’ paper, the term “virtual oscillators” is used instead of harmonic oscillators in the classical theory. In fact, they say that “the correspondence principle has led to comparing the reaction of an atom on a field of radiation with the reaction on such a field which, according to the classical theory of electromagnetics, should be expected from a set of ‘virtual’ harmonic oscillators”.

The absorption and emission of radiation are supposed to be coupled with different processes of transition, and thereby with different virtual oscillators. They say that “just this point, however, seems to afford a guidance as regard the way in which the scattering phenomena are related with the activity of the virtual oscillators concerning emission and absorption of radiation”, and that “in a later paper it is hoped to show how on the present view a quantitative theory of dispersion resembling Ladenberg’s theory can be established”. This paper is Kramers’ 1924 paper which we shall consider in Sec. 4, §1 of the present chapter.

About the Compton effect, Bohr-Kramers-Slater take the view that it has been shown “by means of reflection of X-rays from crystals”, and mention
that “Compton has reached a formal interpretation of this effect”. They say that “in contrast to this picture, the scattering of radiation by electrons is, on our view, considered as a continuous phenomenon to which each of the illuminated electrons contributes through the emission of coherent secondary wavelets”.

As has been seen in Sec. 1, §2 of the present chapter, in order to estimate the intensity of the scattered X-ray, Compton assumed that the light quantum would be emitted with a probability proportional to the angular distribution of the Thomson scattering in the rest coordinate system of an electron moving with the “effective velocity”. Bohr-Kramers-Slater call such an electron “an imaginary moving source”, and suppose that “the incident virtual radiation gives rise to a reaction from each electron, similar to that to be expected on the classical theory from an electron moving with a velocity coinciding with that of the above-mentioned imaginary source and performing forced oscillations under the influence of the radiation field”. About this supposition, they add that the fact “that in this case the virtual oscillator moves with a velocity different from that of the illuminated electrons themselves is certainly a feature strikingly unfamiliar to the classical conceptions”. This refers to the difficulty in their supposition that, according to the classical theory of forced oscillations induced by radiation, the electron makes oscillation around its original position as the center. Bohr and the others do not mention that Compton’s “effective velocity” is indeed derived on the basis of the theory of light quanta.

They try to explain away this difficulty by saying that “in view of the fundamental departures from the classical space-time description involved in the very idea of virtual oscillators, it seems at the present state of science hardly justifiable to reject a formal interpretation as that under consideration as inadequate”. “On the contrary, such an interpretation seems unavoidable in order to account for the effects observed, the description of which involves the wave-concept of radiation in an essential way.” Saying so, Bohr-Kramers-Slater negate the particle-nature of radiation in the Compton effect.

§3. Compton-Simon’s experiment (Confirmation of the energy-momentum conservation in an elementary process)

Two experiments were performed in 1925 on Bohr-Kramers-Slater's interpretation of the Compton effect. One was an experiment by Bothe and Geiger on
the coincidence counting of scattered X-ray and recoil electron, and the other was an experiment by Compton and Simon using cloud chamber.

Bothe and Geiger say, in their paper "On the Essence of the Compton Effect; an Experimental Contribution to the Theory of Radiation", that they have made this experiment on their thought that according to the idea of Compton and Debye a recoil electron must be observed simultaneously with each scattered light quantum, while according to the interpretation by Bohr and others the coincidence would be much looser and must vanish in the experimental arrangement practically used.

In this experiment, two point-counters are placed in a hydrogen gas of one atmospheric pressure on two sides of the incident X-ray beam, face to face with each other and at right angles to the beam. One of them counts scattered X-rays and the other recoil electrons, each having an opening of $7 \times 10 \text{ mm}^2$ in size. The incident beam is a beam of continuous X-rays from a 70 kV Coolidge tube of tungsten electrode, passed through an aluminum foil of 1 mm thickness. Suppose $Z$ scattered X-rays and $N$ recoil electrons are counted per second, of which $C$ electrons are in coincidence with X-rays. The coincidence in counting is discriminated with the use of an electric circuit, which determines whether X-rays and recoil electrons are counted within a critical time $t$ or not. If, of $C$ electrons counted in coincidence in this sense, $C_-$ electrons are counted before the X-rays and $C_+$ electrons after the X-rays, the expectation value of the counting in accidental coincidence within a time $\tau$ is given by $E = N Z \tau$, while the probability that the counting in accidental coincidence of more than $C_+$ electrons takes place is given by the Poisson distribution as

$$P = \sum_{C_+}^{\infty} \left( \frac{E^{C_+}}{C_+!} \right) e^{-E}.$$  \hspace{1cm} (43)

(A similar equation holds for $C_-$.)

From two series of measurements with $\tau = 1/100 \text{ sec}$, and one series of measurements with $\tau = 1/1000 \text{ sec}$, they get the total values of $Z = 647$, $C_+ = 49$ and $C_- = 17$. Namely, one recoil electron is counted in coincidence for about ten scattered X-rays. Bothe and Geiger conclude that "this is not understandable according to Bohr's idea, but is to be expected according to

---


59) Counters with a needle-like electrode.
The Interconnection of Wave- and Particle-Natures

the previous idea,\textsuperscript{60} if one takes into account the unavoidable incompleteness of the experimental conditions". The expectation value of the counting in accidental coincidence is found to be $E = 23.4$, and the probability given by Eq. (43) is $P = 2.5 \times 10^{-6}$.

Compton and Simon say, in their paper "Directed Quanta of Scattered X-rays",\textsuperscript{61} that "a study of the scattering of individual X-ray quanta and of the recoil electrons associated with them makes possible, however, what seems to be a crucial test between the two views of the nature of scattered X-rays". According to them, Bothe-Geiger's experiment "affords less definite evidence than does the present one regarding the directive nature of the scattered X-rays". Now, if each scattered quantum is deflected through some definite angle $\phi$ from its incident direction, and the electron which deflects the quantum recoils at an angle $\theta$, one has the relation

$$\tan \frac{\phi}{2} = \frac{-1}{(1 + \alpha) \tan \theta}, \quad (44)$$

where $\alpha = \frac{h}{mc\lambda_0}$, $\lambda_0$ being the wavelength of the incident X-rays. In order to observe the values of $\phi$ and $\theta$, Compton and Simon use an almost monochromatic beam of X-rays, which is obtained by making X-rays from a 140 kV Coolidge tube pass through 6 mm of brass and about 2 mm each of copper and aluminum. The beam is led into a Wilson cloud chamber to be scattered by the air in it. The values of $\theta$ are found by finding the tracks of recoil electrons. In directions with values of $\phi_{\text{obs}}$ near the value $\phi_{\text{calc}}$ calculated for each observed value of $\theta$ through Eq. (44), tracks of secondary electrons are found.

They measure the values of $\theta$ and $\phi_{\text{obs}}$ by observing the negatives, taken with a stereoscopic camera of 5.5 cm focal length, by means of a stereoscope with lenses of the same focal length as those of the camera. The errors in the measurements of $\theta$ and $\phi_{\text{obs}}$ are estimated to be less than $10^\circ$ and $5^\circ$ respectively. 302 plates are obtained in the first series of measurements, and 338 and 511 plates respectively in the second and third series of measurements in which lead shields are used to reduce stray $\beta$-rays. The histogram of the deviation from theoretical angle, $\Delta = |\phi_{\text{obs}} - \phi_{\text{calc}}|$, per $20^\circ$ shows, in the sum of the second and third series, that the number of the events with $\Delta < 20^\circ$ is $(7.9 \pm 0.7)$ times larger than the average number of events over the whole range $(0 \leq \Delta \leq 180^\circ)$.

\textsuperscript{60} The idea due to Compton and Debye.

About half of the events with $\Delta < 20^\circ$ are due to those processes in which there are one or two other recoil electrons that are in no way associated with the secondary $\beta$-ray. Though the origin of the remaining half cannot be assigned definitely, they say that undoubtedly some are due to stray X-rays which could not be completely eliminated and some are probably due to $\beta$-rays of radioactive origin. They thus conclude, that "at least a large part of the scattered X-rays proceed in directed quanta of radiant energy". "These results do not appear to be reconcilable with the view of the statistical production of recoil and photoelectrons proposed by Bohr, Kramers and Slater. They are, on the other hand, in direct support of the view that energy and momentum are conserved during the interaction between radiation and individual electrons."

Slater's paper in 1925 cited above\(^{62}\) was written after the experiments of Bothe and Geiger and of Compton and Simon were performed. In this paper Slater says, concerning his acceptance of the idea due to Bohr and Kramers, that "under their suggestion, I became persuaded that the simplicity of mechanism obtained by rejecting a corpuscular theory more than made up the loss involved in discarding conservation of energy and rational causation". He makes an excuse for this with the words "the changes made in adopting this view are thus not fundamental", and stresses that he is different from Bohr and Kramers by saying "the simplest solution of the radiation problem then seems to be to return to the view of a virtual field to guide corpuscular quanta".

In any way, through the confirmation by Compton-Simon's experiment of the individuality of the Compton effect and the validity of the law of energy-momentum conservation in it, the theory of light quanta became firmly grounded. At the same time, the intermediating viewpoint due to Schrödinger, Duane and Compton of deriving the Doppler effect and the Bragg formula for reflection, with the use of Bohr's frequency relation and Sommerfeld's quantum condition on the basis of the theory of light quanta, got depended in its significance.

\(^{62}\)Ref. 56) in §1 of the present section.
§1. Transition probabilities and the theory of dispersion of light

The trial that was announced in Bohr-Kramers-Slater's paper, that is, the trial to establish "a quantitative theory of dispersion resembling Ladenburg's theory", was done in Kramers's 1924 paper "The Law of Dispersion and Bohr's Theory of Spectra".  

Before this work, Ladenburg attempted to bring transition probabilities into the classical theory of dispersion of light, in a 1921 paper "The Quantum Theoretical Interpretation of the Number of Dispersion Electrons" and in a 1923 paper with the collaboration of Reiche "Absorption, Scattering and Dispersion in Bohr's Theory of Atoms".

Now, in the classical theory, if a light wave of frequency $\nu$ is incident upon a gas, which contains $N_0$ electrons per unit volume each making a harmonic oscillation of frequency $\nu_0$, the electric field of the light wave makes every electron in the gas undergo forced oscillation, giving rise to a change in the electric polarization of the gas. As a result it follows that the refractive index $n$ of the gas obeys the dispersion formula

$$n^2 - 1 = \frac{N_0(e^2/m)}{\pi(\nu_0^2 - \nu^2)}$$

as a function of $\nu$. Ladenburg interprets the number $N_0$ of dispersing electrons in a quantum theoretical way as follows.

If the gas is in thermal equilibrium with radiation, the energy $J_{EI}$ radiated from a unit volume of the gas per unit time is, as in the case of Planck's resonator, given by

$$J_{EI} = \left(\frac{\pi e^2}{m}\right)N_0\rho_0,$$

where $\rho_0$ is the energy density of radiation of frequency $\nu_0$. On the other hand, according to Einstein's theory of thermal radiation, the energy radiated, $J_Q$, and that absorbed, $A_Q$, by a unit volume of the gas per unit time are equal to

---

63) H. A. Kramers, Nature 113 (1924), 673.
64) R. Ladenburg, Zeit. f. Phys. 4 (1921), 451.
65) R. Ladenburg und F. Reiche, Naturwiss. 11 (1923), 584.
each other, and are given by

\[ J_Q = A_Q = N_i \left( \frac{p_k}{p_i} \right) A_k^i \left( \frac{c^3}{8\pi \nu^2} \right) \rho, \]  

(47)

where \( N_i \) is the number of the atoms in the stationary state \( i \), \( A_k^i \) the probability of transition from the state \( k \) to the state \( i \), \( p_k \) the statistical weight of the state \( k \), \( p_i \) that of the state \( i \), \( \nu \) the frequency of the radiation associated with the transition, and \( \rho \) the energy density of the thermal radiation of frequency \( \nu \). Ladenburg puts \( J_{EI} \) equal to \( J_Q \) and expresses \( N_0 \) in terms of \( A_k^i \) by considering that \( \nu_0 \) and \( \rho_0 \) are equal to \( \nu \) and \( \rho \) respectively. Because \( A_k^i \) is independent of the temperature, Ladenburg assumes that the expression so obtained of \( N_0 \) is valid "outside black body cavity", and considers it as "the quantum theoretical interpretation of the dispersion constant \( N_0 \)".

The transition probability \( A_k^i \) "can be calculated approximately from Bohr's correspondence principle in very simple cases" only. Ladenburg attempts to get \( A_k^i \) from the observed values of \( N_0 \) by means of "the quantum mechanical interpretation" mentioned above. For example, from the dispersion of the \( H_\alpha \) - and \( H_\beta \)-lines by hydrogen, the ratio between the values of \( N_0 \) for these lines is known to be \( N_\alpha / N_\beta \sim 4.5 \), from which he estimates the ratio between the transition probabilities for the \( H_\alpha \)- and \( H_\beta \)-lines to be \( A_\alpha / A_\beta \sim 4 \), by taking the statistical factor of the stationary state of principal quantum number \( n \) as \( n(n + 1)/2 \) according to Sommerfeld.

Generalizing the case treated by Ladenburg, Kramers starts in his paper under consideration from the classical expression for the electric dipole moment \( P \), which is induced by the electric field \( E \) of an incident light wave of frequency \( \nu \), in an atom having absorption lines of frequencies \( \nu_1, \nu_2, \ldots \),

\[ P = E \sum_i f_i \frac{e^2 / 4\pi^2 m}{(\nu_i^2 - \nu^2)}, \]

(48)

where \( f_i \) denotes the number of such electrons in an atom that have the natural frequency \( \nu_i \). Kramers considers that, though "the present state of the quantum theory does not allow a rigorous deduction of the laws" concerned with the reaction of an atom to radiation, "it is, however, possible to establish a very simple expression for \( P \)".

\( ^{66} \)Cf. Sec. 3, §2, Chap. 3 of the present volume.
Kramers supposes, thus, that if an atom in a stationary state may perform transitions to states of higher energy by absorbing radiation of frequencies $\nu_i^i(i = 1, 2, \ldots)$, and transitions to states of lower energy by emitting radiation of frequencies $\nu_j^j(j = 1, 2, \ldots)$, $P$ is expressed by

$$P = E \sum_i A_i^a \tau_i^a \frac{e^2}{m 4\pi^2(\nu_i^a - \nu^2)} - E \sum_j A_j^e \tau_j^e \frac{e^2}{m 4\pi^2(\nu_j^e - \nu^2)}.$$  (49)

In this equation, $A_i^a$ is the probability coefficient of the spontaneous transition from the state $i$ to the state under consideration, and $1/\tau_i^a = \frac{8\pi^2 e^2 \nu_i^2}{3mc^3}$ is the classical theoretical decrement per unit time of the energy of an electron performing harmonic oscillation of frequency $\nu_i^a$. $A_j^e$ and $1/\tau_j^e = \frac{8\pi^2 e^2 \nu_j^2}{3mc^3}$ are the corresponding quantities for the spontaneous transition from the state under consideration to the state $j$. The statistical weight is assumed, for the sake of simplicity, to be the same for all the states. When the atom is in the ground state, the second term in the right-hand side of Eq. (49) is not necessitated, and Eq. (49) becomes “equivalent” to Ladenburg’s formula. The underlying consideration of Kramers is to take into account transitions to states of lower energies from a state of large quantum number, for which the transition probability coefficients $A_i^a$ and $A_j^e$ can be calculated by means of the correspondence principle, and to determine the signs of the contributions from these transitions so that the resultant expression of $P$ for such a state becomes the same as the classical one.\(^{67}\)

Kramers says that “the reaction of the atom against the incident radiation can thus formally be compared with the action of a set of virtual harmonic oscillators inside the atom”. If one compares, however, $A_r$ in Eq. (49) with $f$ in Eq. (48), “one might introduce the following terminology; in the final state of the transition the atom acts as a ‘positive virtual oscillator’ of relative strength $+f$; in the initial state it acts as a negative virtual oscillator of strength $-f$”. He considers that “however unfamiliar this ‘negative dispersion’ might appear from the point of view of the classical theory, it may be noted that it exhibits a close analogy with the ‘negative absorption’ which was introduced by Einstein, in order to account for the law of temperature radiation on the basis of the quantum theory”. Namely, the second term in the right-hand side of Eq. (49) is thought to correspond to the ‘negative

\(^{67}\)Cf. descriptions given shortly later.
in-radiation’ in Einstein’s terminology, that is, the induced emission represented by $B_m^p$.  

Breit criticizes, in his paper in 1924 “The Quantum Theory of Radiation”,
that “from the point of view of the virtual oscillators of Bohr, Kramers and Slater, the second negative term of Kramers is somewhat dissatisfactory, because an oscillator would give rise only to a term of the first positive type”.  

Breit points out also that in the case of the linear oscillator, the expression of Kramers at long wave-lengths becomes

$$P = E[n - (n - 1)] \frac{e^2}{m4\pi^2(\nu^2 - \nu^2)} = \frac{Ee^2}{m4\pi^2(\nu^2 - \nu^2)} ,$$

in accordance with the result due to the classical theory, but “the same result may be also derived as”

$$P = E\left(\frac{1}{2} + \frac{1}{2}\right) \frac{e^2}{m^24\pi^2(\nu^2 - \nu^2)} .$$

Breit emphasizes that “the dependence of $P$ on $\nu$ can be derived from a consideration of a ‘virtual orbit’ rather than a virtual oscillator”. About the virtual orbit, he supposes that “its reaction to the external field is comparable with that of an electron, the orbit of which is the mean of the orbits between the two stationary states”.

Kramers, who had been shown Breit’s paper by the editor of Nature before its publication, published a note in conjunction with it. In this note, Kramers states that “it may be emphasized that the notion ‘virtual oscillators’ used in my former letter does not mean the introduction of any additional hypothetical mechanism”, and that “the fundamental importance … is, as mentioned in my former letter, indicated by the necessity, pointed out by Einstein, of introducing the idea of negative absorption”.

Kramers negates, thus, the necessity of virtual oscillators for the theory of dispersion, though Bohr-Kramers-Slater state, as we have seen in the precedent section, that “the correspondence principle has led to comparing the reaction of an atom on a field of radiation with the reaction on such a field which, according

---

68) Note the relation $A_m^p = (8\pi\hbar^3/c^3)B_m^p$ due to Einstein.

69) G. Breit, Nature 114 (1924), 310.

70) For $A_e^p$ and $A_e^e$, the values for a harmonic oscillator obtained from the correspondence principle are used.

71) H. A. Kramers, Nature 114 (1924), 310
to the classical theory of electrodynamics, should be expected from a set of ‘virtual’ harmonic oscillators”. Kramers follows Ladenburg’s idea of connecting the number of dispersing electrons with the transition probability, and fulfills the requirement from the correspondence principle in its original meaning that quantum theoretical treatments should coincide with the respective conclusions of the classical theory, by making use of Einstein’s “negative in-radiation”.

By the way, “the correspondence principle” in the words of Bohr and the others quoted just above has a meaning different from its original one. The correspondence principle in the quoted words means to make some classical theoretic-mechanism correspond to a quantum theoretical process. That is, it means to take some substantialistic model in the stage of seeking for essentialistic laws — no matter whether the model will be successful or not.

Now, Kramers explains in his note under consideration his dispersion formula as follows. Let the multiple Fourier series expansion of the electric dipole moment $M$ of a multiply periodic system of degree of freedom $f$ be

$$M = \sum_{\tau} C_{\tau} \cos 2\pi \omega t; \quad \omega = \tau_1 \omega_1 + \cdots + \tau_f \omega_f,$$

in which $\tau$ is the abbreviation of the set of integers $\tau_1, \tau_2, \ldots \tau_f$. The dipole moment $P$ induced in the system through the forced oscillation by an external electric field $E \cos 2\pi \nu t$ is given by

$$P = \frac{E}{2} \sum_{\tau} \frac{\partial}{\partial J} \left( \frac{C_{\tau}^2}{\omega^2 - \nu^2} \right) \cos 2\pi \nu t,$$

in terms of the set of the action variables $J_1, J_2, \ldots J_f$, which determine states of the system, $\partial/\partial J$ being the abbreviation of $\tau_1 \partial/\partial J_1 + \tau_2 \partial/\partial J_2 + \cdots \tau_f \partial/\partial J_f$.

In the classical theory one has from the Hamiltonian $H$ the relation

$$\omega = \frac{\partial H}{\partial J},$$

but the quantum-theoretical frequency $\nu_q$ is determined by the relation

$$\nu_q = \frac{\Delta H}{\hbar},$$

where $\Delta H$ is the difference in energy between the states for which $J_1, J_2, \ldots J_f$ are different, respectively, by $\tau_1 h, \tau_2 h, \ldots \tau_f h$. Considering that $(2\pi \nu)^4 C_{\tau}^2 / 3c^3$ is the energy radiated per unit time by the system according to the correspondence principle, Kramers assumes that the replacement of $\partial/\partial J$ in Eq. (51)
with a difference similar to that for $\nu_q$ results in the two terms of Kramers' dispersion formula given in Eq. (49).

Considerations similar to Kramers' are also seen in Born's paper "Quantum Mechanics" received about five weeks before the above cited note of Kramers. Born derives a dispersion formula of Kramers type, by developing a classical-theoretical perturbation treatment of the forced oscillation induced in a multiply periodic system by an external field, and by replacing the derivative appearing in the resultant formula, which is equivalent to Eq. (51), with the difference. But, Born gives no particular discussion on the relation of the constants in the two terms thus obtained to the transition probabilities. He acknowledges Bohr for discussions on "the contents of this paper", but states that "our line of reasoning will be independent of ... the frame work of Bohr-Kramers-Slater's theory, such as the statistical interpretation of the transfer of energy and momentum". Bohr's paper is not cited in the note of Kramers.

§2. The theory of light quanta and the dispersion theory

Kramers generalizes his dispersion theory, in the paper "On the Dispersion of Radiation by Atoms" written in collaboration with Heisenberg five months later than the note cited above, to include scatterings of light incoherent with the incident one. Kramers and Heisenberg express the oscillating electric field $E(t)$ of an incident monochromatic plane wave of light as

$$E(t) = R(Ee^{2\pi i\nu t}),$$  \hspace{1cm} (52)

where $R$ denotes the real part of the quantity following it. Electrons in an atom undergo forced oscillations by this electric field, giving rise to an oscillating electric dipole moment

$$P(t) = R(Pe^{2\pi i\nu t})$$  \hspace{1cm} (53)

induced in the atom, on account of which spherical electromagnetic waves are emitted from the atom. $P$ in this expression is the same as that treated in Kramers' theory of dispersion.

---

72) M. Born, Zeit. f. Phys. 26 (1924), 379.
Kramers-Heisenberg generalize Eq. (53) to assume the expression

\[
P(t) = R \left[ Pe^{2\pi i vt} + \sum_k P_k e^{2\pi i (\nu + \nu_k) t} + \sum_l P_l e^{2\pi i (\nu - \nu_l) t} \right].
\]  

(54)

Namely, they suppose that the atom emits spherical waves which have frequencies different from that of the incident wave, and therefore are incoherent with the incident wave. They say that “the wave-theoretical statement (Ansatz) Eq. (54) corresponds exactly to Smekal’s result”. 74)

In a paper in 1923 “On the Quantum Theory of Dispersion”, 75) Smekal criticized Darwin’s theory of dispersion which was based on the wave theory. Darwin supposed, in a series of papers, 76) that “when an atom is struck by a wave, there is a certain chance that the atom should emit a secondary wave”, on the assumption that “the wave theory gives a correct account of events outside matter”. The chance is assumed to be proportional to the time derivative of the electric field of the incident wave, and the proportional coefficient, as well as the frequency of the wave divergent from the atom, is supposed to depend on the “ways in which the atom may be excited”. He emphasized that the superposition of fields of radiation emitted by the atom at various instants would give a result similar to that in the classical theory of dispersion, if the frequency of the wave divergent from the atom was equal to the natural frequency in the classical theory.

Smekal states in his paper under consideration that “it is an essential progress that instead of electromagnetic-mechanical eigen-frequencies of the molecular kind in the classical theory, their quantum frequencies are contained” in Darwin’s theory, “in spite of many difficulties”. Smekal considers, contrary to Darwin, that the scattering of monochromatic light is due to the absorption and re-emission of light quanta by atoms, in accordance with Einstein’s idea of the emission and absorption of radiation by atoms in his derivation of the Planck formula of thermal radiation. Namely, he assumes that, if an atom moving in a direction with a velocity \( u \) is in a stationary state of energy \( E_m \), it makes the transition, under the influence of radiation of frequency \( \nu \), to a state of energy \( E_n \) by emitting radiation of a different frequency \( \nu' \), with a probability per unit time \( C_{m,n,\nu,\nu'}^{n,\nu'} \). By this transition, the velocity of the atom

74) The original number of Eq. (54) is (6).
75) A. Smekal, Naturwiss. 11 (1923), 873.
changes from \( u \) to \( u' \), on account of the momenta of the absorbed and emitted radiations, whose magnitudes are \( h\nu/c \) and \( h\nu'/c \) respectively.

Denoting the mass of the atom by \( M \), one has from the conservation of energy

\[
\frac{Mu^2}{2} + E_m + h\nu = \frac{Mu'^2}{2} + E_n + h\nu'.
\]

Smekal call the case of \( m = n \) "normal scattering". Even in this case, the situation \( \nu \neq \nu' \) is possible because of the change of momentum of the atom. Such as case is "called simply translational quantum transition". The case of \( m \neq n \) is called "anomalous scattering". In anomalous scattering, the case of \( \nu = 0 \) or \( \nu' = 0 \) "becomes respectively absorption or emission by in-radiation (Einstein)", for which

\[
h\nu + \frac{Mu^2}{2} + E_m = \frac{Mu'^2}{2} + E_n, \quad (E_m < E_n)
\]
\[
\frac{Mu^2}{2} + E_m = \frac{Mu'^2}{2} + E_n + h\nu. \quad (E_m > E_n)
\]

The case of \( E_m < E_n \) "gives the Doppler effect as Schrödinger showed".

Smekal says that "as is clear from the close connection between scattering and dispersion, one must expect, by suitable use of the quantity \( C_{m,n}^{\nu,\nu'} \), to be able to arrive at a quantum-mechanical interpretation of dispersion". To this standpoint, "a strong support is given by the results of the scattering of sufficiently short X-rays, in particular, by the established frequency change of sideway-scattered secondary radiation, out of which the successful theory of A. H. Compton and P. Debye result". Smekal points out also that to the problem of the reflection of light "in a somewhat special way W. Duane has already drawn his attention". Smekal states that "it is however perhaps still a long way to destroy the dogma of the indispensability of wave-theoretical thoughts in the optics of reflection and interference".

Now, Kramers and Heisenberg consider Eq. (54) as "wave-theoretical statement (Ansatz)" that "corresponds just to the result of Smekal", who has grasped on the basis of the theory of light quanta the process of scattering of light as a process of absorption and emission of light quanta. They note that if we "neglect the small change in velocity of the atom", and "denote the change in energy of the atom by \( h\nu_k \) or \( h\nu_l \), according to whether the change
is in positive or negative direction”, then “the frequency of the light quantum emitted in this process will be clearly given by $v + v_k$ or $v - v_l$”.

Though Kramers-Heisenberg’s paper was written before the experiments of Bothe-Geiger and of Compton-Simon, they state in this paper that “the calculation by means of light quanta has, above all, significance in the fact that it prepares us to bring the macroscopic conservation laws of energy and momentum into connection with the quantum-theoretical ideas in a simple and informative way”, avoiding to comment on the conflict between Bohr-Kramers-Slater’s theory of statistical conservation of energy-momentum and the theory of light quanta. They add that “by the nature of things such considerations do not allow us, however, to draw any conclusion on some corpuscular structure of light, while we must always require that the results so obtained can be brought into agreement with the wave-theoretical description of the optical phenomena in an uncontradictory way”. That is, they attempt to arrive at the wave-theoretical understanding of Smekal’s result obtained on the basis of the theory of light quanta.

Kramers-Heisenberg calculate the reaction of a multiply periodic system irradiated by light, using a method outlined in Kramers’ note cited in the preceding subsection. The system under consideration is described in terms of the action variables $J_1, J_2, \ldots, J_s$ and angle variables $w_1, w_2, \ldots, w_s$, and its energy $H$ is assumed to depend only on the action variables. The fundamental frequencies of the system are given by $\omega_k = \frac{\partial H}{\partial J_k}(k = 1, 2, \ldots, s)$. The electric dipole moment $M(t)$ of the system is developed in multiple Fourier series

$$M(t) = \sum_{\tau_1 \ldots \tau_s} \frac{1}{2} C_{\tau_1 \ldots \tau_s} e^{2\pi i(\tau_1 w_1 + \cdots + \tau_s w_s)}. \tag{55}$$

In order to know the reaction of the system to the electric field given by Eq. (52), they introduce the infinitesimal contact transformation

$$J_k^* - J_k = \frac{\partial K}{\partial w_k^*}, \quad w_k^* - w_k = -\frac{\partial K}{\partial J_k^*},$$

to define a new set of the canonical variables $J_k^*$ and $w_k^*(k = 1, 2, \ldots, s)$. It is found that $K$ defined by

$$K = R \left\{ \sum_{\tau_1 \ldots \tau_s} \left( -\frac{1}{2} \right) \frac{(E \cdot C_{\tau_1 \ldots \tau_s})}{2\pi i(\omega + \nu)} e^{2\pi i(\tau_1 w_1^* + \cdots + \tau_s w_s^* + \nu t)} \right\}$$
gives $J_k^*$ independent of the time $t$ and $w_k^*$ such that $w_k^* = \omega_k^* t$ in the first order approximation of $E$. If the transformed $J_k^*$ and $w_k^*$ are denoted anew by $J_k$ and $w_k$ respectively, $M(t)$ becomes

$$M(t) = M_0(t) + M_1(t),$$  \hspace{1cm} (56)

where $M_0(t)$ has the same form as Eq. (55) with $w_k = \omega_k t$, and $M_1(t)$ is given by

$$M_1(t) = R \left\{ \sum_{\tau_1^0 \cdots \tau_s^0} \sum_{\tau_1 \cdots \tau_s} \frac{1}{4} \left[ \frac{\partial C(E \cdot C')}{\partial J'} \omega' + \nu \right] \right.$$  

$$- C \frac{\partial}{\partial J} \left( \frac{E \cdot C'}{\omega' + \nu} \right) \right\} e^{2\pi i (\omega^0 + \nu) t}. \hspace{1cm} (57)$$

In this equation, $C$ and $C'$ are the abbreviations of $C_{\tau_1 \cdots \tau_s}$ and $C'_{\tau_1, \tau_s}$ respectively and

$$\frac{\partial}{\partial J} = \tau_1 \frac{\partial}{\partial J_1} + \cdots + \tau_s \frac{\partial}{\partial J_s}, \quad \frac{\partial}{\partial J'} = \tau'_1 \frac{\partial}{\partial J_1} + \cdots + \tau'_s \frac{\partial}{\partial J_s},$$

$$\tau_1 + \tau'_1 = \tau_1^0, \ldots, \tau_s + \tau'_s = \tau_s^0,$$

$$\omega' = \tau'_1 \omega_1 + \cdots + \tau'_s \omega_s, \quad \omega^0 = \tau_1^0 \omega_1 + \cdots + \tau_s^0 \omega_s.$$

$M_1(t)$ represents the electric dipole moment induced in the system, corresponding to $P(t)$ given by Eq. (53).

The particular terms with $\omega^0 = 0$ in Eq. (57) give rise to coherent scattered light which has the same frequency as the incident light. These terms are expressed as

$$M_1(t) = R \sum' \frac{1}{4} \frac{\partial}{\partial J} \left[ \frac{C(E \cdot \tilde{C})}{\omega - \nu} + \frac{\tilde{C}(E \cdot C)}{\omega + \nu} \right] e^{2\pi i \nu t} \hspace{1cm} (58)$$

where $\sum'$ means summation over $\tau_1, \ldots, \tau_s$ such that $\omega = \tau_1 \omega_1 + \cdots + \omega_s \tau_s > 0$. In order to get the quantum-theoretical expression for this $M_1(t)$, Kramers-Heisenberg replace the derivative $\partial/\partial J$ with the difference, as in Kramers' note. By this replacement, the classical frequency $\omega = \partial H/\partial J$ becomes the quantum-mechanical frequency $\nu_q = \Delta H/h$. At the same time, because in the limit of large enough quantum number, the probability $\alpha_\nu$ of spontaneous transition of frequency $\nu$ is given, according to the correspondence principle,
by $\alpha_{\nu}h\nu = \{(2\pi\nu)^{4}/3c^{2}\}(A_{\nu} \cdot \vec{A}_{\nu})$ in terms of the amplitude $(1/2)A_{\nu}$ of the component of $M_{0}(t)$ of frequency $\nu$, they suppose that the two terms of positive and negative signs in $M_{1}(t)$ resulting from the replacement of $\partial/\partial J$ with the difference represent the processes of absorption and emission of light.

Thus, for the coherent scattering Kramers-Heisenberg quantum-theoretically put

$$M_{1}(t) = R \sum_{a} \frac{1}{4\hbar} \left\{ \frac{A_{a}(\mathbf{E} \cdot \vec{A}_{a})}{\nu_{a} - \nu} + \frac{\vec{A}_{a}(\mathbf{E} \cdot \vec{A}_{a})}{\nu_{a} + \nu} \right\}
- \sum_{e} \frac{1}{4\hbar} \left\{ \frac{A_{e}(\mathbf{E} \cdot \vec{A}_{e})}{\nu_{e} - \nu} + \frac{\vec{A}_{e}(\mathbf{E} \cdot \vec{A}_{e})}{\nu_{e} + \nu} \right\} e^{2\pi i\nu t}, \tag{59}$$

where $\sum_{a}$ means summation over the frequency $\nu_{a}$ of selected absorption, and $\sum_{b}$ over the frequency $\nu_{e}$ of spontaneous transition. $A_{a}$ and $A_{e}$ denote the quantum-theoretical amplitudes of absorption and emission respectively. If the incident light is linearly polarized so that $A_{a}$ and $A_{e}$ are parallel to $\mathbf{E}$, the strengths $f_{a}$ and $f_{e}$ defined by the relation $f = \alpha_{\nu}\tau_{\nu}$ with $\nu = \nu_{a}$ and $\nu = \nu_{e}$ respectively make Eq. (59) identical with Kramers' dispersion formula. In the mentioned relation, $\tau_{\nu} = 3mc^{3}/8\pi^{2}e^{2}\nu^{2}$ is the lifetime of an electron against losing its energy according to the classical theory.

The incoherent scattering is given by the terms with $\omega^{0} \neq 0$ in Eq. (57). In this case, because $\tau'_{k} \neq -\tau_{k}$ Eq. (57) does not reduce to a single summation as in Eq. (58). If the transition from the state P to the state R is indicated by the $\tau'_{k}$'s and that from the state R to the state Q by $\eta_{k}$'s, there can also be transition from P to S indicated by the $\tau'_{k}$'s and that from S to Q indicated by the $\tau_{k}$'s. The transitions S $\rightarrow$ Q and P $\rightarrow$ R have the same $\tau_{k}$'s but are separated from each other by the $\tau'_{k}$'s, so that they are parallel to each other, so to speak. For this reason, Kramers-Heisenberg put

$$\frac{\partial C}{\partial \nu} \left[ \frac{(\mathbf{E} \cdot C')}{\nu'} \right] \approx \frac{A_{SQ} - A_{PR}}{2} \frac{1}{\hbar} \left[ \frac{(\mathbf{E} \cdot A_{RQ})}{\nu_{RQ} + \nu} + \frac{(\mathbf{E} \cdot A_{PS})}{\nu_{PS} + \nu} \right],$$

where $A_{RQ}$ and $\nu_{RQ}$ are the amplitude and frequency of $R \rightarrow Q$ and the latter is given by $\nu_{RQ} = H(R) - H(Q)$ in terms of the energies $H(R)$ and $H(Q)$ of R and Q, respectively. Similarly, they put

$$C \frac{\partial}{\partial J} \left[ \frac{(\mathbf{E} \cdot C')}{\nu'} \right] \approx \frac{A_{SQ} + A_{PR}}{2} \frac{1}{\hbar} \left[ \frac{(\mathbf{E} \cdot A_{RQ})}{\nu_{RQ} + \nu} - \frac{\mathbf{E} \cdot A_{PS}}{\nu_{PS} + \nu} \right].$$
In this way, Kramers-Heisenberg conclude that Eq. (57) can be expressed quantum-theoretically as

\[ M_i(t) = R \left[ \sum_{Q} \sum_{R} M(Q, P; R) + \sum_{Q} \sum_{R} M(P, Q; R) \right], \quad (60) \]

where

\[ M(P, Q; R) = \frac{1}{4\hbar} \left[ \frac{A_{QR}(E \cdot A_{RP})}{\nu_{RP} + \nu} - \frac{A_{RP}(E \cdot A_{QR})}{\nu_{QR} + \nu} \right] e^{2\pi i(\nu_{QR} + \nu)t}, \quad (61) \]

and the first summation is taken over all \( Q \)'s with \( \nu_{QP} < \nu \), and the second over all \( Q \)'s with \( \nu_{QP} < -\nu \). Kramers-Heisenberg assume that when \( \nu_{QP} + \nu > 0 \) the initial state of the transition between \( P \) and \( Q \) is \( P \), and the final state is \( Q \), while when \( \nu_{QP} + \nu < 0 \) they are in the inverse relation. This is because Kramers-Heisenberg consider that Eq. (61) gives rise to scattered light of frequency \( \nu_{QP} + \nu \), accompanied by a "change of state" of the atom to lose an amount of energy \( h|\nu_{QP} + \nu| \), while the atom gains or lose an amount of energy \( h\nu \) on account of the incident light (Einstein's positive and negative in-radiations). In the case of \( \nu_{QP} + \nu > 0 \), the atom gains the energy \( h\nu \) and loses at the same time an energy, \( h(\nu_{QP} + \nu) \), losing an energy \( h\nu_{QP} \) in total. In the case of \( \nu_{QP} + \nu < 0 \), for which \( \nu_{QP} < 0 \) always, the atom loses an energy \( h\nu \) as well as \(-h(\nu_{QP} + \nu)\), losing an energy \(-h\nu_{QP} \) in total.

Transitions of the former kind are the ones that were considered by Smekal. Kramers-Heisenberg states that the "transitions of the latter kind hardly arise so naturally from the standpoint of light quanta". It may be noted that, seen from Smekal's "anomalous scattering" based on the theory of light quanta, \( \nu'(=\nu_{QP} + \nu) \) with positive or negative \( \nu_{PQ} \) should be positive, while Kramers-Heisenberg's wave-theoretical Eq. (54) includes pairs of components oscillating as \( e^{\pm2i\nu't} \), as is seen from the notation for the real part \( R \), so that it includes essentially \( \nu' \) of negative sign. This is related to the fact that a light wave is determined by its intensity and phase.

In this way, Kramers and Heisenberg generalized further, by adopting Smekal's idea of the change in frequency of scattered light based on the theory of light quanta, the dispersion formula due to Kramers to include incoherent scattering of light. Kramers' dispersion formula was, in turn, obtained by generalizing the result due to Ladenburg, who attempted to connect transition probabilities with the classical dispersion formula based on the wave theory.
Thus, in the course of investigating the interconnection between the wave and particle theories of light, the electric dipole moment of an atom came to be considered to consist of elements corresponding to its various transitions \([A_{QR}, \text{ etc. in Eq. (61)}]\). In the case of the hydrogen atom, the \(x\)-component of its electric dipole moment is given by \(-ex\). If \(x\) consists of elements of such kind, in what manner given are the elements of \(x^2\). Heisenberg treated this problem, arriving at the so called matrix mechanics.

On the other hand, de Broglie arrived at the concept of phase wave, as we have described in the present chapter, by attributing a rest mass to the light quantum in order to grasp, from the standpoint of the theory of light quanta, the Planck formula of radiation in which the particle- as well as the wave-nature of radiation is included. In this course, too, the interconnection between the dual natures of light was investigated, and material particles came to be considered to possess the wave-nature. The quantum condition turned out to be understandable on the basis of the wave-nature of the electron, and at the same time the diffraction of the electron was predicted. Thus course of investigation was about to reach the finale, in which new developments were revealed. Schrödinger found the equation governing the wave of the electron, arriving at the so-called wave mechanics. Wave mechanics and matrix mechanics became unified to bring about quantum mechanics. The analysis of this process will be given in Vol. III of our book.
I have been able to accomplish Vol. II by having Prof. Masayuki Nagasaki of Rikkyo University as the co-author. He was a colleague when I were at Rikkyo University as a professor, and is my coworker up to now. He made efforts furthermore to write detailed explanatory notes to my following books (each being published by Keiso Pub. Co.),

Collected Works of M. Taketani 3, "War and Science",

Collected Theories of the Present Age by M. Taketani 2, "The Nuclear Age". Each of these two sets of notes is excellent even if it is considered as an independent paper.

He contributed also to my book

"Social Responsibility of Scientists — On Nuclear Weapons" with the following paper

"On Einstein’s Letter to the Journal ‘Report’", which wiped off completely a wrong popular saying about Einstein with respect to the atomic bomb.

He wrote further

"Explanatory Notes" to the republished edition by photocomposition of

"The Formation and Logic of Quantum Mechanics, Vol. I", of which mention is made in the preface to the present volume. These notes are excellent and include critique of some popular mistakes.

Prof. Nagasaki worked as the co-author with might and main. He read and analyzed in detail a great number of papers, extracting essential sentences in them to work out our descriptions by means of words used in original papers. He kept in contact with me to discuss the composition of the present Vol. II.
We have been able to accomplish it thus through our collaboration. To him I would like to express my sincere thanks.

Dec. 14, 1990

Mituo Taketani
THE FORMATION AND LOGIC OF QUANTUM MECHANICS

Vol. III
The Establishment and Logic of Quantum Mechanics
This page is intentionally left blank
THE FORMATION AND LOGIC OF QUANTUM MECHANICS

Vol. III
The Establishment and Logic of Quantum Mechanics

by
MITUO TAKETANI
MASAYUKI NAGASAKI

Translation from the Japanese Edition by Masayuki Nagasaki
Preface to Volume III

We wrote out in the summer of 1990 and published in the spring of 1991 "The Formation and Logic of Quantum Mechanics Vol. II — The Way to Quantum Mechanics" as a collaboration of Taketani and Nagasaki, which succeeds to "The Formation and Logic of Quantum Mechanics Vol. I — The Formation of Atomic Models" by Taketani. Now, we are today in a position to tell the reader that we have completed our work of writing "The Formation and Logic of Quantum Mechanics Vol. III — The Establishment and Logic of Quantum Mechanics". We would like again to beg pardon from the reader of his long wait since the publication of Vol. I.

In Vol. II, we analyzed the process just before the proposal of matrix mechanics and wave mechanics. Firstly, in Chaps. 1 and 2, we traced the development of logic in the following progress: the Planck constant \( h \), which had been introduced phenomenologically for abstract oscillators in the theory of thermal radiation, was clarified to be a constant governing the thermal vibration of atoms in a solid body as well as characterizing the antagonism between undulatory and corpuscular natures; trials were made to make clear atomic phenomena with the use of \( h \), and to systematize the combination rule of spectral lines by comparing it with the result of mathematical analog models; Nicholson proposed in his theory of spectral lines based on an atomic model of Nagaoka type the quantum nature of angular momentum, by which the footing to go to the Bohr model of atoms was set up; and so on.

Then in Chap. 3, we treated the development of the so-called "old quantum theory", in which the quantum nature of angular momentum established by the Bohr model was generalized as the quantum condition by Sommerfeld,
and with the use of it the existence of stationary states of an atom and the validity of the frequency relation of spectral lines were confirmed. In that treatment, we made clear those contradictions that the quantum condition was not to be applicable with necessity to the relativistic motion of an electron, and the spectral terms derived from the quantum condition gave rise to superfluous spectral lines if it was used as itself together with the combination rule, namely, the contradictions which the quantum condition had as an article of the lawfulness in a substantialistic stage.

Also in Chap. 3, we attached importance to the fact, that Einstein introduced on the analogy of the phenomenon of $\gamma$-decay the transition probability for transitions between stationary states, to derive the Planck formula of radiation, and stressed further by treating the recoil of an electron in its emission and absorption of radiation the importance of constructing a truly quantum theoretical theory of radiation. Though Bohr set forth in order to calculate the transition probability the correspondence principle, from which the selection rule of spectral lines was derived with a provisional success, it turned out that the value of transition probability was very much different whether it was calculated about the initial or final state. It was one of the themes of Chap. 3 to make clear the role played by the correspondence principle and the contradiction contained in it.

In Chap. 4, we traced the substantialistic process, in which the clarification of the structure of atoms was attempted, by making use of the results of the spectroscopy of X-rays as well as light and the summary of phenomenological knowledge in chemistry on the basis of the quantum condition and selection rule. There we treated also the process, in which the half-integer quantum number and the anomalous coefficient 2 of magnetic dipole moment were introduced from the analysis of the anomalous Zeeman effect, and this brought about the difficulty of the coefficient 2, together with the facts that in Stern-Gerlach's experiment on a beam of atoms of silver the beam split only into two, though this experiment was taken to give an evidence of the quantization of direction, and that the results of experiments on the magnetization of matter required a half value for the specific charge of an electron, if the magnetization was supposed to be due to the rotational motion of electrons.

We analyzed further the substantialistic process, in which the fine structure of spectral terms different from the relativistic fine structure coming from the quantum condition was found to exist through the study of the spectrum of X-rays, and about it Landé noticed, in contrast to Bohr who introduced
two kinds of azimuthal quantum numbers, the analogy to the doublet term in optical spectra, giving rise to the proposal of the exclusion principle, the introduction of electron's spin, and the substitution of the rotation of an electron round itself in a nonrelativistic orbit for the movement of the perihelion of its relativistic orbit.

In the last Chap. 5, we took from the viewpoint of the interconnection of undulatory and corpuscle natures the process, in which the Doppler effect, the scattering of X-rays by an electron (the Compton effect) and their reflection by a crystal (the Bragg reflection), namely, what had been considered characteristic undulatory phenomena, came to be derived from the standpoint of the theory of light quantum through the mediation of the conservation law of energy and momentum and the quantum condition, and also the process in which de Broglie's wave theory of matter resulted from his supposition of a finite mass about the light quantum in his attempt to formulate the theory of light quantum.

In connection with this, we attached importance to the fact, that Bohr together with Kramers and Slater denied the conservation law of energy and momentum for the Compton effect and proposed the hypothesis of their statistical conservation, which was immediately negated experimentally. We traced further the process in which Kramers and Heisenberg, who tried to understand the dispersion of light from the standpoint of the undulatory theory by applying the transition probability, derived the dispersion formula, by introducing transition amplitudes corresponding to various transitions of an atom, on the adoption of Smekal's idea that the scattering of light by an atom should be due to the absorption and emission by it of a light quantum.

In this way, through the complex processes, of which we have mentioned an outline above, the substantialistic systematization was carried out of the structure of atoms and the interaction of an atom with radiation, and the interconnection between the undulatory and corpuscular natures was inquired into, so that the proposal of matrix mechanics and wave mechanics was prepared.

Now, in the present Vol. III, we shall finally treat the proposal of matrix mechanics and wave mechanics, analyze the process of the establishment of quantum mechanics through their mutual penetration and unification, and make clear the stereo-structural logic contained in quantum mechanics thus accomplished. We say the unification of matrix mechanics and wave mechanics, but it is never such one as is often described in textbooks of quantum mechanics. It is said, for example, that both were different entirely in forms but
identical completely in contents, or that both were unified because it became clear that both represented one and the same content in different mathematical forms. These viewpoints are static and of projected view, and do not grasp the antagonism between the concepts in matrix mechanics and wave mechanics.

As we shall see in Chap. 1, matrix mechanics started from Heisenberg's attempt to get some rules of calculation about the "quantum theoretical amplitudes" which Kramers and Heisenberg introduced for an inner atomic electron. The rules of calculation thus found were formulated by Born and Jordan with the use of matrix, and by Dirac with the use of $q$-algebra, and the commutation relations and the so-called Heisenberg equation of motion were discovered.

Matrix mechanics was, in view of its start, a phenomenological theory based on corpuscular theory. It depended on "old quantum theory", which had been obtained in the general process that started from the an sich and phenomenological theories of spectral lines and of thermal radiation, and came to the essentialistic theory of quantum mechanics. "Old quantum theory" had restrictions, as the für sich theory in a substantialistic stage.

The quantum condition, which was one of the lawfulness in "old quantum theory", premised the existence of quantum numbers. de Broglie's wave theory of matter gave prospects of deducing quantum numbers with necessity, not from the promise but from natural physical boundary conditions, like the number of nodes in the vibration of a string. The prospects were realized by Schrödinger's setting up the wave equation for de Broglie's wave, and Schrödinger proceeded then to the proposal of wave mechanics. This process was not even as is described in present-day textbooks of quantum mechanics.

In Chap. 2, the zigzag way, along which Schrödinger walked struggling to construct an entirely new mechanics, will be made clear through the analysis of his original papers. There, we shall analyze the points that have rarely been taken until now, such as the ones that Schrödinger treated gas, just before his first paper on wave mechanics, as an aggregate of proper vibrations of de Broglie's wave, that he related at first the energy to the square of frequency because he thought the so-called Schrödinger equation not to be a wave equation but to be a frequency equation, that in order to overcome this difficulty he made investigations from various viewpoints, and that after having made many trials to grasp the meaning of wave function $\psi$ he came to grasp $|\psi|^2$ as the weight function by discarding the meaning of the wave as an intuitive real being.
As we shall see also in Chap. 2, Schrödinger criticized Heisenberg’s viewpoint as Machism, in which the relations among the “observable” quantum theoretical amplitudes, each being supposed to give a transition probability when absolutely squared, were supposed only to be dealt with in the new mechanics, and showed that every matrix element was given by proper solutions of the Schrödinger equation. Wave mechanics was thus a substantialistic theory based on undulatory theory. Wave mechanics as such was in antagonism to matrix mechanics which was a phenomenological theory based on corpuscular theory, and both interpenetrated and became unified to give rise to quantum mechanics as the \emph{an und für sich} law in an essentialistic stage.

The three stages in the development of the cognition in physics, namely, phenomenological, substantialistic and essentialistic stages, “do not proceed uniformly in the sense that ‘how it becomes then and then’, but proceed along a series of coiling turns of these three stages” (M. Taketani, “On the Formation of the Newton Mechanics”, Kagaku, Aug. 1942). As we have mentioned above, also in the process of transition from the substantialistic to essentialistic stage in the general development of quantum mechanics, the antagonism between phenomenological and substantialistic theories, together with that between corpuscular and undulatory theories, was existent making a small coiling turn, so to speak, and was developing.

“The three stages should not be taken in a mechanical way, because the logic of each stage appears in various forms in accordance with the reality”, as is noted in Taketani’s paper cited just above. “The three stages are the fundamental distinguishing features for the analysis of the development of physics. Two of them may be tangled in some case, and three in another case.” (M. Taketani, “Modern Physics and Epistemology”, Sizen Kagaku, Jul. 1946).

In the present Vol. III, we shall trace the real development of this logic, in accordance with the ideas and expressions in original papers as in Vol. II. We have mentioned above that the two in antagonism interpenetrated and were unified, but the interpenetration does not take place without mediation. In Chap. 3, the logic of the interpenetration of matrix mechanics and wave mechanics through mediation and the logic of their unification will be analyzed in accordance with the real development, that Born treated the scattering of particles, which could not be dealt with in matrix mechanics, by means of wave mechanics, grasping the wave function as the amplitude of the existence probability of a particle, Dirac found matrix mechanics, in which observable quantities only were considered calculable, to be unable to determine the matrix
element to give the transition probability for a system of electrons, seeking for
the solution to this difficulty in the symmetry property of the wave function for
the permutation of electrons, and so on. This kind of analysis has seldom been
done in the books hitherto published on the history of quantum mechanics.

In Chap. 4 we shall elucidate the stereo-structural logic that is characteristic
of quantum mechanics unified as the law in essentialistic stage. There we
shall analyze the fact that Heisenberg took the uncertainty relation, which
is one of the phenomenal forms of the law of quantum mechanics, to be a
principle, proposing the negation of causality, Bohr introduced on the basis
of Heisenberg's idea the term of complementality, which does not represent
a philosophical nor scientific concept, advocating the so-called "Copenhagen
spirit", and so on.

We shall further analyze the logical structure of v. Neumann's theory of
observation based on "Copenhagen spirit", and show that the interpretation
by these three of quantum mechanics and observation problem is of projected
view lacking in the understanding of the stereo-structural logic of quantum
mechanics. It will be treated also in Chap. 4 that, though Schrödinger's fa-
mous example of cat was a sharp attack on the contradiction in theories of
observation of v. Neumann's style, Schrödinger himself was not able to go out
completely from his substantialistic logic, in spite of his struggle to grasp the
logic of quantum mechanics. We shall analyze, besides, the controversy be-
tween Einstein-Podolsky-Rosen and Bohr, and some wrong interpretations of
the observation problem.

From these analyses, the reader will keenly feel the importance of grasping
enough that quantum mechanics has its characteristic stereo-structural logic,
and the core of the quantum mechanical observation consists in the fact based
on the stereo-structural logic of quantum mechanics, that the objective bound-
dary between the object and observer lies at the observing apparatus in which
an irreversible change takes place.

We hope that the table of chronological style put at the beginning of the
present volume will serve the reader, in deepening his understanding of the
papers treated in Vol. III from the viewpoint of the interpenetration between
matrix mechanics and wave mechanics.

Up to the completion of our book with the present Vol. III, we have been
given words of encouragement and expectation by many persons, to whom we
are very thankful. Our thanks are also due to Miss Masako Ishii of Keiso Pub. Co. for her lending aid as in the case of Vol. II.

Mituo TAKETANI
Masayuki NAGASAKI
September 30, 1992
This page is intentionally left blank
Relations Among Papers Contributed to the Establishment of Quantum Mechanics

The dates shown are those of the reception of each paper. When they were not given, the dates of issue are shown. Each arrow indicates citation. Upward arrows indicate citations in advance. The numbers in parentheses mean Chap.-Sec., Subsec.
Relations among Papers contributed to the Establishment of Quantum Mechanics

The dates shown are those of the reception of each paper. When they were not given, the dates of issue are shown. Each arrow indicates citation.
Upward arrows indicate citations in advance. The numbers in parentheses mean Chap.-Sec., Subsec..

1925

7.29 Heisenberg: Derivation according to the corresponding principle of the rules of calculation of the quantum theoretical amplitudes of an electron introduced phenomenologically in Kramers-Heisenberg's theory of dispersion of light. (1-1, §1—§4)

9.27 Born-Jordan: Formulation with the use of matrices of Heisenberg's rule of calculation. Derivation of the commutation relation and the equation of motion. (1-2, §1—§3)

11.7 Dirac: Formulation with the use of q-algebra by paying attention to the similarity of Heisenberg's rule of calculation to Poisson's parenthesis equation. Derivation of the commutation relation and equation of motion for a many body system. (1-3, §1, §2)

11.16 Born-Heisenberg-Jordan: Generalization of matrix mechanics to many body systems. Formulation of canonical transformation and perturbation method. Derivation of the quantum number of angular momentum, selection rule and relative intensity of the normal Zeeman effect, quantum fluctuation of the energy of a string. (1-2, §4—§6)

12.22 Lanczos: Attempt to relate the discreteness in Matrix mechanics to the discrete proper solutions of an integral equation of a continuous quantity. (1-4, §1)

12.15 Schrödinger: Theory of gases as aggregation of the proper oscillations of the de Broglie wave. Derivation of Einstein's formula of quantum fluctuation. (2-1, §4)

1926

1.5 Born-Wiener: Attempt with the introduction of operators to solve the difficulty of matrix mechanics being non-applicable to aperiodic phenomena. (1-4, §2)

1.17 Pauli: Derivation of the energy levels of hydrogen atom by means of matrix mechanics. (1-2, §7)

1.22 Dirac: Attempt to derive the energy levels of hydrogen atom by a q-algebraic treatment of the equation of orbit of the electron. (1-3, §3)
3.16 Heisenberg-Jordan: Calculation of the anomalous Zeeman effect of hydrogen atom by means of matrix mechanics on the assumption of Uhlenbeck-Goudsmit's model (1925.10.17) and the anomalous coefficient of magnetic moment. (1-2, §5)

3.27 Dirac: Application of $g$-algebra to angular momentum. Treatment of the angle variable canonically conjugate to angular momentum by $g$-algebra. (With a note to mention that Born-Heisenberg-Jordan obtained independently the formula of angular momentum.) (1-3, §3)

4.29 Dirac: Attempt to generalize matrix mechanics relativistically. Discussion of the Compton effect. (3-3, §2)

6.11 Heisenberg: Discussion of the perturbation method for a system of identical atoms having resonant degeneration. Expression of the obtained result in terms of the permutation symmetry of $W$. In spite of his thought that wave mechanics was not truly wave theoretical because of the use of the interaction between particles given by classical theory. (3-1, §1, §2)

1.27 Schrödinger: The first paper. Derivation of the energy levels of hydrogen atom by setting up a frequency equation for the de Broglie wave in the atom. The energy was supposed to be related to the frequency squared. (2-1, §2, §3)

2.23 Schrödinger: The second paper. Notice of the surface of $W$ (Hamilton's principal function) = $\text{const}$, moving geometric optical way in the configuration space. Setting up of the wave equation on the assumption of $W$ = the phase of the wave in the configuration space. Application of it to diatomic molecules etc. Comparison of the quantum condition with the theory of Fresnel bands. (2-2, §1-§3)

3.18 Schrödinger: The paper on the relation'. Verification of the fact that matrix elements can be given by the eigenfunctions of the wave equation. Assertion of the equality of wave and matrix mechanics on the assumption of the time dependence of matrix elements given by matrix mechanics. Critique of matrix mechanics as Maclaurian. (2-3, §1-§3)

5.10 Schrödinger: The third paper. Application of perturbation method to the wave equation. Calculation of the energy shift in the Stark effect of hydrogen atom, by getting the electric dipole moment on the assumption that the charge density $= \omega \psi / \partial t$. Comparison of the result obtained with the experimental data using the interpretation of matrix elements in matrix mechanics. (2-4, §1, §2)

6.18 Wentzel: Derivation of the quantum condition with the use of the so-called WKB approximation. Calculation of the second order Stark effect. (2-2, §3)
8.26 Dirac: Indication of the inability of matrix mechanics to determine matrix elements of a many body system. Solution of this difficulty by the permutation symmetry of $\Psi$, with the introduction of the F-D as well as B-E statistics. Setting up of the first order differential equation of $\Psi$, the $m$-th coefficient of expansion of $\Psi$ in the eigen functions, to discuss the absorption and induced emission of light on the assumption of $\langle \omega \rangle$ being the number of atoms in the $m$-th state. (3-1, §3–§6)

10.16 Born: Supposition of $|\omega(0)|^2$ to be the probability of the system being in the $m$-th state, with the use of the time dependent Schrödinger equation and the interpretation in matrix mechanics to relate the matrix element to transition probability. Derivation of the formula of transition probability for the absorption and induced emission of light, by the treatment of light as a group of pulses over whose time duration the average was taken. (3-2, §1)

11.6 Heisenberg: Supposition of the square of a matrix element of unitary transformation $S_m$ to be a probability, from the energy $(E + E')/2$ of a system of two interacting identical atoms of energies $E$, $E'$ and by analogy with the matrix elements $\pm 1/\sqrt{2}$ of unitary transformation in this case. (3-2, §3)

11.19 Wentzel: Derivation of the formula of Rutherford scattering with the use of Born's theory of scattering. Calculation of photo-electric effect. (3-3, §1)

12.2 Dirac: Generalization of matrix mechanics with the use of $S$-function to make it able to include continuous dynamical quantities, for the purpose of treating aperiodic phenomena such as scattering. Introduction of the transformation function between discrete and continuous representations, to show that Schrödinger's wave function is a transformation function. Derivation of the same result as Born's for the scattering of particles, by the adoption of Heisenberg's idea because of the nature of the transformation function of being a generalized matrix element of unitary transformation. (3-2, §3, §4)

6.21 Schrödinger: The fourth paper. Supposition of $\Psi$ to be essentially complex for the purpose of eliminating the energy from the wave equation. Setting up of the partial differential equation of first order about the time. Derivation from it of the conservation law of $|\Psi|^2$. Interpretation of $|\Psi|^2$ as 'weight function'. (2-4, §3, §4; 5, §1)

6.25 Born: Application of wave mechanics to the scattering of a particle, an aperiodic motion. Interpretation of $|\Psi|^2$ as the probability of existence of the particle by the introduction of corpuscular theoretic viewpoint. (2-5, §2, §3)

9.19 London: Application of the method of canonical transformation in matrix mechanics to the wave equation, with the result that $\Psi$ is a state quantity to determine a vector in the Hilbert space. (3-2, §2)

9.29 Gordon: Relativistic wave equation. Calculation of the Compton effect on the supposition of the wave field to be the source of classical electromagnetic field. (3-3, §2)
1928

1.2 Dirac: Presentation of the relativistic wave equation (the Dirac equation) by the "relativistic generalization" of the Pauli spin matrices. (3-3, §7)

1.26 Jordan-Wigner: Formulation of the treatment of a system of particles obeying the F-D statistics by the imposition of a "special non-commutable property", that is the anti-commutation relation, on the expansion coefficients of the de Broglie wave into its proper oscillations. (3-3, §6)

2.2 Dirac: Generalization of the Dirac equation to the case of the existence of electromagnetic field. (3-3, §7)

10.30 Klein-Nishina: Derivation of the Klein-Nishina formula of the Compton effect, by means of the calculation of the electromagnetic field caused by the electron wave given by the Dirac equation in the case of the existence of an incident electromagnetic wave. (3-3, §7)

1929

3.19 Heisenberg-Pauli: Quantization of the Dirac relativistic electron field and electromagnetic field, by the imposition of the anti-commutation relation and commutation relation, respectively, on their values in small cells taken as the canonical variables. (3-3, §6)
## Contents

Preface to Volume III                      v

Relations Among Papers Contributed to the Establishment of
Quantum Mechanics                       xiii

Volume III. The Establishment and Logic of
Quantum Mechanics                        1

Chapter 1. The Proposal and Formulation of Matrix Mechanics 3

1. Heisenberg’s “Quantum Condition” . . . . . . . . . . . . . . . . . . . . . . . . . . 3
   §1. Heisenberg’s thought of quantum theory . . . . . . . . . . . . . . . . . . . . . 3
   §2. The “aggregate” of quantum-theoretical amplitudes and
       the combination rule of amplitudes .......................... 7
   §3. New “quantum-theoretical quantum condition” ................. 10
   §4. Application of the new quantum condition to oscillator
       and rotator .................................................. 14

2. Born-Jordan’s Formulation with Matrices ................. 18
   §1. From “aggregates” of amplitudes to matrices .......... 18
   §2. Presentation of the essentially complex commutation
       relation ...................................................... 21
   §3. The laws of time variation and of energy conservation
       in matrix mechanics ........................................ 23
   §4. Generalization by Born-Heisenberg-Jordan to systems
       of multiple degrees of freedom ............................. 25
3. Dirac's Formulation by Quantum Algebra ........................................... 42
§1. Poisson bracket expressions and commutation relations ............... 42
§2. q-numbers and c-numbers — quantum algebra .......................... 46
§3. Application of the quantum algebra ................................................... 49
§3.1 Electron orbit in hydrogen atom ......................................................... 49
§3.2 Angular momentum and its canonical conjugate coordinate .................. 52
4. Attempts at the Interpretation of Matrix Mechanics ....................... 55
§1. Lanczos's attempt by means of integral equation ......................... 55
§2. Born-Wiener's attempt by means of operators .......................... 60

Chapter 2. From the Proposal of Wave Mechanics to Quantum Mechanics 67
1. Schrödinger's Wave Equation ............................................................. 67
§1. Theory of gas based on de Broglie's wave theory ......................... 67
§2. The wave equation and quantum numbers ...................................... 72
§3. The frequency of standing de Broglie wave and that of radiation .................. 76
2. "Geometric Mechanics" and "Undulatory Mechanics" .................... 79
§1. The Fermat principle and the Hamilton principle .......................... 79
§2. Assumption of the wave frequency and the wave equation .......... 84
§3. Quantum condition, theory of Fresnel zones and WKB approximation .................................................. 88
3. Relation between Wave Mechanics and Matrix Mechanics ............ 93
§1. Wave functions and matrix elements .............................................. 93
§2. Mathematical equivalence between matrix mechanics and wave mechanics ........................................... 96
§3. Phenomenology and substantialism ............................................... 99
4. Perturbation Theory and Dispersion Formula of Light ................ 104
§1. Perturbation theory for energy eigenvalues .......................... 104
§2. Intensity of spectral lines ................................................................. 108
§3. Time dependence of wave function and complex wave function .................................................. 113
§4. Time-dependent perturbation and dispersion formula of light .................................................. 116

5. The Physical Meaning of Wave Function .......................................................... 120
§1. $\psi\overline{\psi}$ as "the weight function" and the continuity equation ........................................... 120
§2. Born's particle-theoretical interpretation of wave function ............................................... 124
§3. Coefficient in eigenfunction expansion and the frequency of occurrence of a state ........... 128
§4. Wave equation and scattering probability of a particle ...................................................... 131

Chapter 3. The Establishment of Quantum Mechanics 137
1. Interpenetration between Matrix- and Wave-Mechanics .................................................... 137
§1. Matrix- and wave-mechanics for many-body systems ................................................... 137
§2. Heisenberg's analysis of the "resonance" between atoms .............................................. 141
§3. Dirac's analysis of a system of identical particles ............................................................ 146
§4. Bose-Einstein statistics and Fermi-Dirac statistics ......................................................... 150
§5. Time as a canonical variable and the equation of motion .................................................. 154
§6. Dirac's consideration of Einstein's theory of transition probability .................................. 156

2. The Unification as Quantum Mechanics .............................................................................. 160
§1. Time variation of $\psi$ and transition probabilities .............................................................. 160
§2. Wave function as a "state quantity" in Hilbert space ....................................................... 165
§3. The concept of probability for matrix elements of transformation ................................... 169
§4. Matrices with continuous row and column and the $\delta$-function ........................................ 173
§5. Wave function as a transformation function .................................................................... 177
§6. Formulation based on the concept of probability amplitude ............................................ 182
§7. The unified formulation with the use of abstract Hilbert space ......................................... 189

3. Developments in Various Problems: Spin Wave Functions, Symmetry for Particle Exchange, Relativistic Formulation 197
§1. Applications of Born's statistical interpretation ................................................................. 197
§1.1 Rutherford scattering ....................................................................................................... 197
§1.2 Photoelectric effect ......................................................................................................... 199
§2. Relativistic formulation of the theory and the Compton effect ................................................. 201
§2.1 Relativistic formulation of the Schrödinger equation ......................................................... 201
§2.2 Compton effect ...................................................................................................................... 204
§3. Unification of light-quantum theory and electromagnetic wave theory ..................................... 206
§3.1 Bose-Einstein statistics and the second quantization ............................................................. 207
§3.2 Application to the system of light quanta ................................................................................. 210
§4. Introduction of spin wave functions ............................................................................................ 213
§4.1 Darwin’s de Broglie wave of vector type ................................................................................ 213
§4.2 Pauli’s wave functions for spin ............................................................................................. 214
§5. Symmetries of wave function and atomic and molecular structures ........................................ 217
§5.1 Symmetries of wave function and spectral multiplets ............................................................ 217
§5.2 Symmetric properties of the wave function of diatomic molecule and the nuclear spin ....... 219
§5.3 Symmetry property of wave function and the binding of hydrogen molecule ......................... 222
§6. Formulations of the quantization of wave fields ......................................................................... 223
§7. The Dirac equation and the Klein-Nishina formula .................................................................... 228

Chapter 4. The Logic of Quantum Mechanics ................................................................................. 235
1. Uncertainty “Principle” and Complementarity “Principle” ......................................................... 235
§1. Heisenberg’s uncertainty relation .............................................................................................. 235
§2. The uncertainty relation and commutation relation .............................................................. 238
§3. Heisenberg’s interpretation of quantum-mechanical laws ....................................................... 241
§4. The uncertainty relation and Bohr’s idea of complementarity-reciprocity .............................. 246
§5. “Copenhagen spirit” and its limit ............................................................................................. 248
2. v. Neumann’s Theory of Observation and Critique of It ............................................................ 253
§1. v. Neumann’s statistical operator and “mixed” aggregate ...................................................... 253
§1.1 Statistical Operator .................................................................................................................. 253
§1.2 “Uniform” aggregate and “mixed” aggregate ....................................................................... 257
§2. v. Neumann’s theory of observation .......................................................................................... 259
§3. Critique by Schrödinger ........................................ 264
§3.1 The case of a cat .............................................. 264
§3.2 The $\psi$ function as "the catalog of expectation" .... 267
§4. Critique by Taketani ............................................ 269
§4.1 The stereo-structural logic of quantum mechanics ... 269
§4.2 Irreversible macroscopic process — the objective
boundary between observer and object ....................... 273

3. Around the Logic of Quantum Mechanics .................. 276
§1. Einstein-Podolsky-Rosen's doubt .......................... 276
§2. Bohr's 'reply' to EPR ............................................. 280
§3. Furry's 'support' to Bohr ...................................... 283
§4. Irreversible processes .......................................... 286
§4.1 Green's model .................................................... 286
§4.2 About the so-called "negative measurement" ..... 290
§5. "Each photon interferes only with itself" — Dirac's
expression ............................................................ 293

Index for Volumes I, II, III ................................. 299
Volume III

The Establishment and Logic of Quantum Mechanics
This page is intentionally left blank
Chapter 1

The Proposal and Formulation of Matrix Mechanics

1. Heisenberg's "Quantum Condition"

§1. Heisenberg's thought of quantum theory

During the period of about one year from the middle of 1925 to the middle of 1926, quantum theory made brilliant developments, bringing forth the establishment of quantum mechanics in 1927. These developments took place in the finale, that was reached after the intricate course in which, as has been described in Vol. II "The Way to Quantum Mechanics" of the present book, the various contradictions appearing one after another were studied, systematized and clarified. The process of development advanced with matrix mechanics based on Heisenberg's theory proposed in 1925, and with wave mechanics proposed by Schrödinger in 1926, as the main themes. The former is treated in the present chapter, and the latter will be discussed in the next chapter.

Heisenberg presented, in his paper received on July 29, 1925 “On Quantum Mechanical Reinterpretation of Kinematic and Mechanical Relations”,\(^1\) "new quantum-mechanical relations" by introducing new interpretations of Ritz's combination rule and the quantum condition. Heisenberg arrived at his quantum-mechanical relations on the basis of Kramers' theory of dispersion of light completed about a year earlier and Kramers-Heisenberg's generalization of this theory made about half a year earlier.

As we have seen in Secs. 3 and 4, Chap. 5, Vol. II, Bohr did not fully believe the theory of light quanta, and was particularly in doubt about the conservation of energy and momentum in each individual process of emission

\(^1\)W. Heisenberg, Zeit. f. Phys. 33 (1925), 879.
or absorption of a light quantum. When Slater\textsuperscript{2}) attempted in 1924 to grasp
the interconnection between the wave and particle natures of radiation with
the idea that distant atoms might communicate with each other by means of a
virtual field of radiation, which would guide light quanta, Bohr and Kramers
pointed out to Slater that "the advantages of this essential feature would be
kept, although rejecting the corpuscular theory, by using the field to induce a
probability of transition rather than by guiding corpuscular quanta".\textsuperscript{3})

Bohr, Kramers and Slater\textsuperscript{4)\textsuperscript{5}) negated the conservation of energy and momentu
in each individual process of radiation, proposing the statistical con-
servation of energy and momentum. They supposed that a "consistent way
of describing the interaction between radiation and atoms" could be obtained
by assuming an interaction, in the meaning of the correspondence principle,
of the virtual field with virtual oscillators. Their theory of statistical con-
servation of energy and momentum was, however, immediately negated by
Compton-Simon's experiment.\textsuperscript{5})

On the other hand, Kramers\textsuperscript{6)\textsuperscript{7}) stressed that "the notion 'virtual oscilla-
tors' does not mean the introduction of an additional hypothetical mechanism”.
He generalized Ladenberg's result\textsuperscript{8}) which related the relative number of elec-
trons having a specified frequency in the dispersion formula of light according
to the classical wave theory to the transition probability, by the introduction
of Einstein's negative absorption (induced emission). Kramers gave a disper-
sion formula involving the difference between two kinds of terms, which he
interpreted as arising from the displacement of the differential in the classical
theory with the difference. Kramers and Heisenberg\textsuperscript{9}) generalized further the
formulation due to Kramers by taking into account Smekal's idea\textsuperscript{10}) of consid-
ering the scattering of light as arising from the absorption and emission of light

\textsuperscript{2}) J. C. Slater, Nature \textbf{113} (1924), 307 [the same as Ref. 55) in Chap. 5, Vol. II].
\textsuperscript{3}) J. C. Slater, Nature \textbf{116} (1925), 278 [the same as Ref. 56) in Chap. 5, Vol. II].
\textsuperscript{4}) N. Bohr, H. A. Kramers and J. C. Slater, Phil. Mag. \textbf{47} (1924), 785 [the same as
Ref. 57) in Chap. 5, Vol. II].
\textsuperscript{5}) A. H. Compton and A. W. Simon, Phys. Rev. \textbf{26} (1925), 289 [the same as Ref. 61) in
Chap. 5, Vol. II].
\textsuperscript{6}) H. A. Kramers, Nature \textbf{113} (1924), 673 [the same as Ref. 63) in Chap. 5, Vol. II].
\textsuperscript{7}) H. A. Kramers, Nature \textbf{114} (1924), 310 [the same as Ref. 71) in Chap. 5, Vol. II].
\textsuperscript{8}) R. Ladenburg, Zeit. f. Phys. \textbf{4} (1921), 451; R. Ladenberg und F. Reiche, Naturwiss.
\textbf{11} (1923), 584 [the same as Refs. 64) and 65) in Chap. 5, Vol. II].
\textsuperscript{9}) H. A. Kramers und W. Heisenberg, Zeit. f. Phys. \textbf{31} (1925), 681 [the same as Ref. 73)
in Chap. 5, Vol. II].
\textsuperscript{10}) A. Smekal, Naturwiss. \textbf{11} (1923), 873 [the same as Ref. 75) in Chap. 5, Vol. II].
quanta, to include incoherent scattering of light. Thus, in the course of investigating the interconnection between the wave and particle theories of light, the electric dipole moment of an atom came to be considered as consisting of amplitudes of frequencies corresponding to the various transitions.

Now, Heisenberg starts his paper under consideration\(^1\) with the words — "As is well known, against the formal rules which are generally used in quantum theory for the calculation of observable quantities (the energy in hydrogen atom, for example), a serious objection is raised that these rules of calculation contain as essential constituents relations between quantities which cannot in principle be observed (such as the position and period of revolution of an electron, for example"

Heisenberg says that “the hope that these quantities, unobservable at present, could be made available experimentally \ldots could be seen as justified, if the mentioned rules be consistent in themselves and applicable to a certain unlimited domain of quantum-theoretical problems”.

Heisenberg considers, however, that the experience shows that “only hydrogen atom and its Stark effect fit into these formal rules of quantum theory”, “the reaction of an atom on a periodically changing field certainly cannot be described by the mentioned rules”, and “an extension of the quantum rules to the treatment of atoms with many electrons has finally proved to be impossible”.

This consideration of Heisenberg reflects the fact that the selection rules, transition probabilities and so on, which were derived on the basis of the quantum condition and the correspondence principle, were confronted with various difficulties, as we have seen in detail, because of the restrictions characteristic of the quantum condition and the correspondence principle as the lawfulness in a substantialistic stage, together with the lack of the knowledge concerning the substantial structure of objects, such as the spin of the electron. In the case of Kramers’ calculation of the intensity of the spectral lines of the Stark effect of hydrogen atom,\(^11\) it was necessary to take some kind of average over the initial and final states. We should not forget, however, the processes through which such difficulties and contradictions were investigated and clarified with various attempts and efforts.

Heisenberg discusses then, as follows — “It is usual to describe this failure of the quantum-theoretical rules, which are essentially characterized by the application of classical mechanics, as deviation from the classical mechanics. This description can however hardly be seen as meaningful, if one considers

\(^1\)H. A. Kramers, Dan. Vid. Selsk. Skrifter. naturvidmat. Afd. 8 (1919), III. 3 [the same as Ref. 61] in Chap. 3, Vol. II.
that the (quite generally valid) Einstein-Bohr frequency condition represents
so complete a refusal of the classical mechanics, or, better, from the standpoint
of wave theory, of the kinematics on which this mechanics bases, and that in
the simplest quantum-theoretical problems any validity of classical mechanics
cannot simply be intended for”.

“In this situation it seems more advisable to give up completely the hope
of observing the hitherto unobservable quantities (like the position and period
of revolution of an electron), and therefore to admit at the same time that the
partial agreements of the mentioned quantum rules with the experience are
more or less accidental, and then to attempt to develop a quantum-theoretical
mechanics analogous to the classical mechanics, in which only relations between
observable quantities appear”. Heisenberg thus proceeds to formulate “such a
new quantum-theoretical mechanics”, by considering “as the most important
first statements, Kramers’ theory of dispersion and the works developed further
on the basis of it”.

As we have described in Sec. 4, §2, Chap. 5, Vol. II, Kramers-Heisenberg
propose for the quantum-theoretical expression for the electric dipole moment
$M_1(t)$, induced by the electric field $E(t) = R(EE^{2\pi ivt})$ of the incident wave
of light, of an atom in a stationary state $P$, which gives rise to the incoherent
scattering of light, the following formula,

$$M_1(t) = R \left\{ \sum_Q \sum_R M(Q, P; R) + \sum_Q \sum_R M(P, Q; R) \right\}.$$ 

In this formula, $R$ means the real part and $M(Q, P; R)$ is defined by

$$M(Q, P; R) = \frac{1}{4\hbar} \left\{ \frac{A_{QR}(E \cdot A_{RP})}{\nu_{RP} + \nu} - \frac{A_{RP}(E \cdot A_{QR})}{\nu_{QR} + \nu} \right\} e^{2\pi i(\nu_{QP} + \nu)t},$$

where $A_{PQ}$ is the amplitude of the transition $P \to Q$, and $\nu_{PQ}$ is given by
$\hbar \nu_{PQ} = H(P) - H(Q)$ in terms of the energies $H(P)$ and $H(Q)$ in the states
$P$ and $Q$ respectively, $A_{QR}$, $\nu_{QR}$, etc. being similarly defined. $M(P, Q; R)$ is
obtained by interchanging $P$ and $Q$ in $M(Q, P; R)$.$^{12}$

Kramers-Heisenberg supposed that, if the unperturbed electric dipole mo­
ment $M_0(t)$ was expressed as a Fourier series $M_1(t) = R \sum_q (1/2) A_q e^{2\pi iv_q t}$

$^{12}$)The atom in $P$ transits to various $Q$ via various $R$, giving rise to waves of incoherent
light of frequency $\nu_{QP} + \nu$. The first summation in the expression of $M_1(t)$ is taken over
the $Q$’s with $\nu_{QP} < \nu$, and the second over the $Q$’s with $-\nu_{QP} (= \nu_{PQ}) > \nu$. 

The Proposal and Formulation of Matrix Mechanics
According to the classical theory, the transition probability of the transition accompanying the emission of radiation of frequency of \( \nu_q \) was given by 
\[
(2\pi)^4 \nu_q^2 |A_q|^2 / 3hc^3
\]
from the correspondence principle. They considered that \( A_{PQ} \) etc. were "the amplitudes characteristic of quantum-theoretical transitions", which should substitute the classical-theoretical amplitudes \( A_q \)'s. On the basis of this consideration, Heisenberg assumes now that every physical quantity of an atom consists of quantum-theoretical components with frequencies corresponding to various transitions.

§2. The "aggregate" of quantum-theoretical amplitudes and the combination rule of amplitudes

According to the classical theory, the strength of the electric or magnetic field of radiation due to a moving electron is, in the approximation of dipole radiation, of the order of \( eiu/rc^2 \), where \( u \) is the speed of the electron and \( r \) is the distance from the electron. In the approximation of quadruple radiation it is of the order of \( eiu/rc^3 \), and terms of the order of \( euu^2/rc^4 \) appear in higher approximations. "The higher approximations can be simply calculated if the motion of the electron, or its Fourier expansion is given". Heisenberg asks "how these higher terms must appear in quantum theory", and says that "one will expect similar things in quantum theory". He stresses his viewpoint that "this question has nothing to do with electrodynamics, but is, this seeming particularly important, of purely kinematic nature".

Heisenberg puts the question in its "simplest form", and asks "if a quantum-theoretical quantity appearing in place of the classical quantity \( x(t) \) is given, what quantum-theoretical quantity will appear in place of \( x^2(t) \)?". He says that, in order to answer this question, "we should remember that it was not possible in quantum theory to attribute to the electron a point in space as a function of time by means of observable quantities".

"The emission of radiation is however attributed to the electron also in quantum theory", and "this radiation is in the first place described by the frequencies". They appear as functions of two variables, \( n \) and \( \alpha \), in the quantum-theoretical form

\[
\nu(n, n - \alpha) = (W(n) - W(n - \alpha))/h,
\]
(1)

instead of the classical form

\[
\nu(n, \alpha) = \alpha \nu(n) = (\alpha/h) dW/dn,
\]
(2)
where "\( nh = J \), \( J \) being one of the canonical constants". That is, in the classical theory, the action variable \( J \) is a constant of a periodic motion with energy \( W \), in terms of which the frequencies are given by \( \nu = \alpha dW/dJ \). Since Heisenberg expresses \( J \) as \( nh \), \( \nu \) becomes a function of \( n \) and \( \alpha, n \) being a continuous variable in this respect.\(^{13}\)

For these frequencies, "one can write down the combination relations".

Classical:

\[
\nu(n, \alpha) + \nu(n, \beta) = \nu(n, \alpha + \beta),
\]

Quantum-theoretical:

\[
\nu(n, n - \alpha) + \nu(n - \alpha, n - \alpha - \beta) = \nu(n, n - \alpha - \beta).
\]

For the description of the radiation, "in the second place the amplitudes are necessary". The amplitudes can be regarded as complex vectors, and determine "polarization and phase". The amplitudes are also functions of "the two variables \( n \) and \( \alpha \)" and, with the angular frequency \( \omega = 2\pi \nu \), are represented by the following expressions:

Classical:

\[
R\{A_\alpha(n)e^{i\omega(n)\alpha t}\};
\]

Quantum-theoretical:

\[
R\{A(n, n - \alpha)e^{i\omega(n,n-\alpha)t}\}.
\]

Heisenberg says that "at first any physical meaning does not seem to belong to the phase (contained in \( A \)) in quantum theory, as the frequencies in quantum theory are not in general commensurable with their harmonic oscillations".

Heisenberg supposes, however, that "phase has a definite meaning also in quantum theory" by considering as follows. Let a quantity \( x(t) \) in the classical theory be represented as\(^{14}\)

\[
x(t) = \sum_{\alpha=-\infty}^{+\infty} A_\alpha(n)e^{i\omega(n)\alpha t}.
\]

---

\(^{13}\) \(ad/dJ\) here corresponds to \( \tau_1 \partial/\partial J_1 \) in Sec. 4, §2, Chap. 5, Vol. II, and \( W(n) - W(n - \alpha) \) corresponds to \( \Delta H \) in the same place.

\(^{14}\) For an \textit{a periodic motion}, Heisenberg writes a similar expression in which the summation is replaced by the integration with respect to \( \alpha \). The same procedure is taken in the following equations.
If \( x^2(t) \) is expressed similarly as

\[
x^2(t) = \sum_{\beta = -\infty}^{+\infty} B_\beta(n) e^{i\omega(n)\beta t}, \tag{8}
\]

in the classical theory \( B_\beta(n) \) is given by

\[
B_\beta(n) e^{i\omega(n)\beta t} = \sum_{\alpha = -\infty}^{+\infty} A_\alpha(n) A_{\beta - \alpha}(n) e^{i\omega(n)(\alpha + \beta - \alpha)t}. \tag{9}
\]

This equation noted down by Heisenberg satisfies of course the classical combination rule of frequencies Eq. (3).

For the corresponding quantum-theoretical quantity, "such a combination does not seem possible without arbitrariness because of the equal rights of \( n \) and \( n - \alpha \), and for that reason does not seem meaningful". Heisenberg considers, however, that "one can regard the aggregate of the quantities \( A_{n, n - \alpha}(n) e^{i\omega(n, n - \alpha)t} \) as the representative of the quantity \( x(t) \) and then try to answer the question: How is the quantity \( x^2(t) \) represented?". This means that, while Eqs. (3) and (9) in the classical theory result from the existence of harmonic oscillations and are therefore concerned with only one fundamental oscillation specified by \( n \), Eq. (4) representing Rydberg-Ritz's combination rule are concerned with three independent states specified by \( n, n - \alpha \) and \( n - (\alpha + \beta) \).

Heisenberg thus takes as the assumption about the quantum-theoretical relation to substitute for Eq. (9) the equation

\[
B(n, n - \beta) e^{i\omega(n, n - \beta)t} = \sum_{\alpha = -\infty}^{+\infty} A(n, n - \alpha) A(n - \alpha, n - \beta) e^{i\omega(n, n - \beta)t}, \tag{10}
\]

which "seems the simplest and most natural assumption" and "arises almost inevitably from the combination rule of frequencies".\(^{15}\) With this equation "one recognizes that the phase of the quantum-theoretical \( A \) has a great physical meaning". In a similar manner, one can find the quantum-theoretical analogue of \( x^3(t) \), and of \( x^n(t) \), and furthermore of a function \( f(x(t)) \) of \( x(t) \) if it may be expanded in a power series in \( x(t) \). Also, if two quantities \( x(t) \) and \( y(t) \) are characterized by the amplitudes \( A \) and \( B \) respectively, the amplitude \( C \) characterizing the product \( x(t) y(t) \) is given by

\(^{15}\)In this equation, Heisenberg uses \( \beta \) instead of \( \alpha + \beta \) in Eq. (4).
Classical:

\[ C_\alpha(n) = \sum_{\alpha = -\infty}^{+\infty} A_\alpha B_{\beta-\alpha}(n), \]  

(11)

Quantum-theoretical:

\[ C(n, n - \beta) = \sum_{\alpha = -\infty}^{+\infty} A(n, n - \alpha) B(n - \alpha, n - \beta). \]

(12)

"Whereas classically \( x(t)y(t) \) is always equal to \( y(t)x(t) \), in the quantum theory this need not in general be the case". As the time derivative of a quantity such as \( u^2(t)/2 \), in quantum theory \( (\dot{u}\dot{u} + uu)/2 \) should be used in place of \( uu \).

§3. New "quantum-theoretical quantum condition"

Having generalized the combination rule of frequencies from the viewpoint of "the kinematics of quantum theory" to get the combination rule of quantum-theoretical amplitudes, which are attributed to a physical quantity correspondingly to various transitions, Heisenberg turns to the "mechanical problem" of determining \( A, \nu \) and \( W \), and re-interprets the quantum condition. He expresses a periodic motion in the classical theory as\(^{16}\)

\[ x = \sum_{\alpha = -\infty}^{+\infty} a_\alpha(n) e^{i\omega_n t}, \]

(13)

and writes down the usual quantum condition (in the so-called old quantum mechanics), in which "one has put the phase integral equal to an integer multiple of \( h \), that is equal to \( nh \)"; as

\[ \oint p dq = \oint m\dot{x} dx = J(= nh). \]

(14)

With the use of the Fourier series development of \( x(t) \), Eq. (13), and the relation \( a_{-\alpha}(n) = \bar{a}_\alpha(n) \),\(^{17}\) this condition is written as

\[ \oint m\dot{x} dx = \oint m\dot{x}^2 dt = 2\pi m \sum_{\alpha = -\infty}^{+\infty} |a_\alpha(n)|^2 \alpha^2 \omega_n. \]

(15)

\(^{16}\)Heisenberg uses here \( a_\alpha(n) \) and \( \omega_n \) in place of \( A_\alpha(n) \) and \( \omega(n) \) respectively in Eq. (7).

\(^{17}\)\( \bar{a}_\alpha(n) \) denotes the complex conjugate of \( a_\alpha(n) \).
Heisenberg says that "such a condition not only fits very forcibly into the mechanical calculation, but also seems arbitrary from the previous standpoint in the sense of the correspondence principle". This is "because correspondingly $J$ is fixed only up to an additive constant as an integer multiple of $\hbar$". He thus considers that in place of Eq. (14),

$$\frac{d}{dn} (nh) = \frac{d}{dn} \int m \dot{x}^2 dt$$

(16)

"would naturally have to enter". That is, the additive constant is eliminated by the differentiation. Inserting this into Eq. (15), he gets

$$h = 2\pi m \sum_{\alpha=-\infty}^{+\infty} \alpha \frac{d}{dn} (\alpha \omega_n |a_\alpha|^2).$$

(17)

Now, Heisenberg states that "if we ask for a quantum-theoretical relation between observable quantities corresponding to Eqs. (15) and (17), the missing uniqueness is automatically recovered", and "in fact only Eq. (17) possesses just a simple quantum-theoretical transformation tied on to Kramers' theory of dispersion", namely

$$h = 4\pi m \sum_{\alpha=0}^{+\infty} \{ |\alpha(n, n + \alpha)|^2 \omega(n, n + \alpha) - |\alpha(n, n - \alpha)|^2 \omega(n, n - \alpha) \}.$$

(18)

This means that by replacing the differential in the reformulated quantum condition given as Eq. (17) with the difference, it is possible, as we shall see shortly later, to get Eq. (18) which derives from Kramers' dispersion formula.

According to Kramers' and Kramers-Heisenberg's theories of dispersion, which we have treated in Sec. 4, §2, Chap. 5, Vol. II, the electric dipole moment $M_1(t)$ of the electron, which is induced by the electric field $E(t) = R(E e^{2\pi i \nu t})$,

---

18) Since the quantum condition demands $n$ to be an integer, the differentiation with respect to $n$ taken by Heisenberg may seem at first glance to be strange. This differentiation could, however, be said to be "natural" at least for large enough $n$ in the sense of the correspondence principle.

19) The original numbers of Eqs. (15) and (17) are (14) and (15), respectively.

20) The original sentence in Heisenberg's paper of "in fact only Eq. (17) possesses just a simple transformation" is "Zwar besitzt eben nur Gleichung (17) eine ... einfache ... Verwandlung". B. L. van der Waerden translates this sentence, in his book "Sources of Quantum mechanics" (North-Holland Co., Amsterdam, 1967), into "We have to admit that ... only equation (17) has a simple ... reformulation". It should be noted, however, that here Heisenberg does not further introduce another assumption.
of the incident wave of light oscillating in the direction of the x-axis, giving rise to the incoherent scattering is given by

\[ M_1(t) = 2e^2 R \sum_{\alpha=0}^{+\infty} \frac{d}{dJ} \frac{|a_\alpha(n)|^2 \omega_n}{(\alpha \nu_n)^2 - \nu^2} E e^{2\pi i \nu t} \]  

(19)

from Eq. (13), since the unperturbed electric dipole moment of the electron is \( M_0(t) = -ex(t) \) in this case.\(^{21}\) Heisenberg replaces the differential in Eq. (19) with the difference, by arguing that the quantum-theoretical frequency given in Eq. (1) can be derived from the classical frequency given in Eq. (2) by replacing the differential in it with the difference, to get the quantum-theoretical expression for \( M_1(t) \) as

\[ M_1(t) = \frac{2e^2}{h} R \sum_{\alpha=0}^{+\infty} \left\{ \frac{|a(n,n+\alpha)|^2 \nu(n,n+\alpha)}{\nu^2(n,n+\alpha) - \nu^2} - \frac{|a(n,n-\alpha)|^2 \nu(n,n-\alpha)}{\nu^2(n,n-\alpha) - \nu^2} \right\} E e^{2\pi i \nu t}. \]  

(20)

This gives Kramers' dispersion formula in the case under consideration.\(^{22}\)

Of this expression, Heisenberg takes the limit that the frequency \( \nu \) of the incident wave of light is much larger than the characteristic frequencies of the dispersing system. Because \( \nu \gg \nu(n,n+\alpha) \) in this limit, Eq. (20) becomes

\[ M_1(t) = -(2e^2/h \nu^2) \sum_{\alpha=0}^{+\infty} \{ |a(n,n+\alpha)|^2 \nu(n,n+\alpha) \} E \cos 2\pi \nu t. \]  

(21)

He considers that in this limit "just as in the classical theory \cdots the oscillating electron behaves like a free electron"\(^{23}\) For a free electron, one has \( M_1(t) = -(e^2/4\pi^2 \nu^2 m) E \cos 2\pi \nu t \), from which Eq. (18) deduces by comparison with Eq. (21). The "simple and quantum-theoretical transformation" from Eq. (17)

\(^{21}\)Cf. Eqs. (55), (56) and (58) in Sec. 4, §2 Chap. 5, Vol. II. \( \alpha d/dJ \) in Eq. (19) above corresponds to \( \tau_1 \partial/\partial \tau_1 \) in Eq. (58) now referred to.

\(^{22}\)Cf. Eq. (59) in Sec. 4, §2, Chap. 5, Vol. II. In Eq. (20) above, \( a(n,n-\alpha) \) has a value only for limited values of \( \alpha \), because there exists the ground state.

\(^{23}\)Since light is represented by an external oscillating electric field acting on the electron in Kramers' as well as Kramers-Heisenberg's theory of dispersion, the Compton effect arising from the particle nature of light is not treated in their theory.
to Eq. (18) can be effected by changing the summation in Eq. (17) to \( \sum_{a=0}^{+\infty} \), in consideration of the factor \( \alpha^2 \) in the summand and the relation \( \alpha_\alpha = \bar{a}_{-\alpha} \) and then replacing \( \alpha d/dn \) by the difference with the method mentioned above.

By the way, a relation equivalent to Eq. (18) is pointed out, shortly before Heisenberg, by Kuhn in his paper “On the total Intensity of Absorption Lines emanating from a given State”. \(^{24}\) Kuhn expresses the amplitude \( M_1 \) of \( M_1(t) \) of hydrogen atom in its ground state as

\[
M_1 = E \left( \frac{e^2}{4\pi^2 m} \right) \sum_{i=0}^{r} \frac{f_i}{\nu_i^2 - \nu^2},
\]

where \( \nu_i \) is the frequency of the \( i \)th absorption line, and \( f_i \) the relative number of the electrons concerned with the \( i \)th line. \(^{25}\) \( f_i \) is related to the transition from the ground state to the \( i \)th state. For \( \nu \gg \nu_i, M_1 \) tends to

\[
M_1 = -E \left( \frac{e^2}{4\pi^2 m \nu^2} \right) \sum_{i=1}^{r} f_i.
\]

If \( \sum f_i \) is set equal to unity, Eq. (23) gives the strength of the Thomson scattering. Kuhn, following Kramers, suggests that when the atom is in an excited state so that it can transit to a state of lower energy, the corresponding \( f_i \) takes a negative value. He then assumes the rule that the sum of the negative as well as of the positive \( f_i \) should be equal to unity. If \( f_i \) is expressed in terms of the transition possibility, this sum rule would give a relation equivalent to Eq. (18). Thomas\(^{26}\) presents a similar theory, too.

Now, Heisenberg regards Eq. (18) as the “quantum-theoretical quantum condition” to be substituted for the usual quantum condition, and attempts to solve the equation of motion of the form

\[
m\ddot{x}(t) + F(t) = 0,
\]

by means of this new condition. It is Eq. (16), obtained through a transformation into differential form, as a “natural” procedure in the sense of the correspondence principle, of Eq. (14) representing the quantum condition in the old quantum mechanics, that acts as the stepping stone for Heisenberg’s jump to the “quantum-theoretical quantum condition”. Heisenberg’s jump is

\(^{24}\) W. Kuhn, Zeit. f. Phys. 33 (1925), 408.

\(^{25}\) Cf. Eq. (48) in Sec. 4, §1, Chap. 5, Vol. II.

\(^{26}\) W. Thomas, Naturwiss. 13 (1925), 627.
attained by noting that Eq. (16), when translated into the form of difference according to the method due to Kramers and Kramers-Heisenberg, becomes identical with the sum rule, which is pointed out by Kuhn and Thomas, and can be derived from Kramers' dispersion formula.

If the equation of motion, Eq. (24), can be solved simultaneously with the condition of Eq. (18), "a complete determination not only of frequencies and energies, but also of quantum-theoretical transition probabilities" is achieved. Heisenberg says that "the actual mathematical performance succeeds, however, for the moment only for simple cases". The simple cases treated by him are the examples of a nonharmonic oscillator and of a rotator.

§4. Application of the new quantum condition to oscillator and rotator

Let the equation of motion of a nonharmonic oscillator be

\[ \ddot{x} + \omega_0^2 x + \lambda x^2 = 0. \] (25)

Heisenberg seeks the solution to this equation by noting that in the classical theory the solution is given in the form

\[ x = \lambda a_0 + a_1 \cos \omega t + \lambda a_2 \cos 2\omega t + \lambda^2 a_3 \cos 3\omega t + \cdots + \lambda^{r-1} a_r \cos r\omega t, \] (26)

where every one of \( a_0, a_1, \ldots, a_r \) is a function of \( \lambda \) to be given as a power series of \( \lambda \) starting from a constant. This form of solution is characteristic of a nonlinear differential equation. Heisenberg tries to get the quantum-theoretical solution by assuming an "analogous" form to represent \( x \) using the following terms.

\[
\begin{align*}
\lambda a(n,n); \ a(n, n-1) \cos \omega(n, n-1)t, \\
\lambda a(n, n-2) \cos \omega(n, n-2)t, \ldots, \\
\lambda^{r-1} a(n, n-r) \cos \omega(n, n-r)t.
\end{align*}
\] (27)

Taking terms up to the first order in \( \lambda \), he derives the following recurrence formulae.\(^{27}\)

\(^{27}\)In the following Eq. (28), Heisenberg writes each \( a_i \) in Eq. (26) as \( a_i(n) \) in accordance with the form of Eq. (13). In the original paper, the first term in the first line of the following Eq. (29) is written as \( \omega_0^2 a_0(n) \), but this should be the misprint of \( \omega_0^2 a_0(n, n) \).
Heisenberg's "Quantum Condition"

Classical:

\[
\begin{align*}
\omega_0^2 a_0(n) + a_1^2(n)/2 &= 0 \\
-\omega^2 + \omega_0^2 &= 0 \\
(-4\omega^2 + \omega_0^2) a_2(n) + a_1^2/2 &= 0 \\
(-9\omega^2 + \omega_0^2) a_3(n) + a_1 a_2 &= 0 \\
\ldots
\end{align*}
\]

Quantum-theoretical:

\[
\begin{align*}
\omega_0^2 a_0(n,n) + \{a^2(n + 1, n) + a^2(n, n - 1)\}/4 &= 0 \\
-\omega^2(n, n - 1) + \omega_0^2 &= 0 \\
\{\omega^2(n, n - 2) + \omega_0^2\} a(n, n - 2) + a(n, n - 1) a(n - 1, n - 2)/2 &= 0 \\
\{\omega^2(n, n - 3) + \omega_0^2\} a(n, n - 3) + a(n, n - 1) a(n - 1, n - 3)/2 + a(n, n - 2) a(n - 2, n - 3)/2 &= 0 \\
\ldots
\end{align*}
\]

On these equations, Heisenberg imposes the quantum condition\(^{28}\)

"Classical (J = nh)"\(^{29}\)

\[
1 = 2\pi m \frac{d}{dJ} \sum_{\tau = -\infty}^{+\infty} \tau^2 \frac{|a_\tau|^2}{4\omega}.
\]

"Quantum-theoretical"

\[
h = \pi m \sum_{\tau = -\infty}^{+\infty} \{ |a(n - \tau, n)|^2 \omega(n + \tau, n) - |a(n, n - \tau)|^2 \omega(n, n - \tau) \}.
\]

In the first approximation both the classical and quantum theories give the same result, that is

\[
a_1^2(n) \quad \text{or} \quad a^2(n, n - 1) = \frac{(n + \text{const})h}{\pi m \omega_0}.
\]

\(^{28}\)The following Eqs. (30) and (31) differ from Eqs. (17) and (18), respectively, by a factor of 4, because in Eqs. (26) and (27) Fourier cosine series is used.

\(^{29}\)The "classical" quantum condition written down by Heisenberg here is not the quantum condition in the old (or classical) quantum mechanics. The latter condition is obtained by integrating Eq. (30) and putting \(J = nh\).
The constant in this equation is, for the case of the quantum theory, determined by the condition \( a(n_0, n_0 - 1) = 0 \) if the ground state corresponds to \( n = n_0 \). For \( n_0 = 0 \), const = 0. From the recurrence formulae in Eq. (28), it is deduced that \( a_T = \kappa(\tau)n^{\tau/2} \), \( \kappa(\tau) \) being a constant independent of \( n \). On the other hand, he gets from Eq. (29) \( a(n, n - \tau) = \kappa(\tau)\{n!/(n - \tau)!\}^{1/2} \), which tends to the result \( a_T \) obtained on the "classical" quantum condition for a large enough \( n \).

Also in the first approximation, Heisenberg derives for the energy \( W = m\dot{x}^2/2 + m\omega_0^2 x^2 + m\lambda x^3/3 \) the following values.

Classical:

\[
W = \frac{nh\omega_0}{2\pi}.
\]  
(32)

Quantum-theoretical:

\[
W = \left(n + \frac{1}{2}\right) \frac{h\omega_0}{2\pi}.
\]  
(33)

In this order of approximation, \( W \) "is constant quantum-theoretically, too".

In order to obtain \( W, a \) and \( \omega \) in higher orders of approximation, Heisenberg takes a simpler example of a nonhormonic oscillator obeying the equation of motion

\[
\ddot{x} + \omega_0^2 x + \lambda x^3 = 0.
\]  
(34)

Classically \( x \) can be put in the form

\[
x = a_1 \cos \omega t + \lambda a_3 \cos 3\omega t + \lambda^3 a_5 \cos 5\omega t + \cdots.
\]  
(35)

Quantum-theoretically, by analogy he assumes the form

\[
a(n, n - 1) \cos \omega(n, n - 1)t, \lambda a(n, n - 3) \cos \omega(n, n - 3)t, \ldots
\]  
(36)

Carrying out calculations in the same way as before, he obtains \( \omega(n, n - 1), a(n, n - 1) \) and \( a(n, n - 3) \) in approximations up to \( \lambda^2 \), from which \( W = m\dot{x}^2/2 + m\omega_0^2 x^2/2 + m\lambda x^4/4 \) is obtained. The results satisfy the relation \( \omega(n, n - 1) = (2\pi/h)\{W(n) - W(n - 1)\} \), where \( W(n) \) is "defined" to be "the constant term" of \( W \), that is, the component \( W(n, n) \) which does not oscillate with time. In the order of approximation under consideration, all the oscillating terms of \( W \), namely the \( W(n, n') \)'s turn out to vanish. Heisenberg says that he "could not prove this generally".
Heisenberg then treats, as an example of rotator, a particle of mass \( \mu \) making a circular motion of radius \( a \). After showing that in this case 
\[
\omega(n, n-1) = n\hbar/2\pi \mu a^2, \quad W(n) = (n^2 + n + 1/2)\hbar^2/8\pi^2 \mu a^2, \quad \omega(n, n-1) = (2\pi/\hbar)[W(n) - W(n-1)],
\]
he considers, in order to discuss the Zeeman effect, the precession of the circular orbit with an angular velocity \( \gamma \) around an external magnetic field acting along the \( z \)-axis. The motion of the particle is now represented, with the use of a quantum number \( m \) for the precession besides the quantum number \( n \) for the circular motion, by

\[
\begin{align*}
z &: a(n, n-1; m, m)\cos\omega(n, n-1)t \\
x + iy &: b(n, n-1; m, m-1)\exp\{i[\omega(n, n-1) + \gamma]t\} \\
x - iy &: b(n, n-1; m-1, m)\exp\{i[-\omega(n, n-1) + \gamma]t\}.
\end{align*}
\]

The equation of motion simply reads \( x^2 + y^2 + z^2 = a^2 \). From the result obtained by applying the combination rule of amplitudes to those given in Eq. (37), together with that due to the “quantum-theoretical quantum condition”

\[
2\pi m\{b^2(n, n-1; m, m-1)\omega(n, n-1)
- b^2(n, n-1; m-1, m)\omega(n, n-1)\} = (m + \text{const})\hbar,
\]

it is deduced that

\[
\begin{align*}
b(n, n-1; m, m-1) &= a \left\{ (n+m)(n+m+1)/4 \left( n + \frac{1}{2} \right) \right\}^{1/2} \\
b(n, n-1; m-1, m) &= a \left\{ (n-m)(n-m+1)/4 \left( n + \frac{1}{2} \right) \right\}^{1/2} \\
a(n, n-1; m, m) &= a \left\{ (n-m)(n+m+1)/ \left( n + \frac{1}{2} \right) \right\}^{1/2}.
\end{align*}
\]

These amplitudes give the intensity of spectral lines in the Zeeman effect. Heisenberg says that “however one cannot simply see that these expressions represent the only solution”. He concludes his paper by saying the following — “Whether a method of determining quantum-theoretical data by relations among observable quantities, like that proposed here, can really be seen to be satisfactory in respect of principle, or whether this method, still more, represents very much a rough attack on the physical, and at first obviously very
involved, problem of quantum-theoretical mechanics, could be recognized, first of all, by a more profound mathematical investigation of the method used here very superficially" —. "A more profound mathematical investigation" came, as we shall see in the next section, to be accomplished in Born-Jordan's work by the use of the theory of matrices. The very importance of Heisenberg's work consists in that it opened the way to the establishment of matrix mechanics.

2. Born-Jordan's Formulation with Matrices

§1. From "aggregates" of amplitudes to matrices

Born and Jordan "developed Heisenberg's approach into a systematic theory of quantum mechanics" in their paper "On Quantum Mechanics"\(^{30}\) received on Sep. 27, 1925. As they write in the abstract of their paper, this was done "with the aid of the mathematical method of matrix". They recognized that the combination rule of quantum amplitudes, Eq. (10) or (12) in the preceding section which Heisenberg assumed, based on hints from the combination rule of frequencies, for amplitudes belonging to a physical quantity correspondingly to various transitions of an atom, was nothing but the multiplication rule of matrices.

Born and Jordan devote the first chapter of their paper under consideration to the presentation of about seven pages of an outline of matrix calculations, citing the books of Bôcher\(^{31}\) and of Courant and Hilbert.\(^{32}\) This may be because matrix calculations were not much familiar. Since the mathematics of matrices is well known nowadays, we omit here the description of their outline. Born-Jordan denote a matrix by a gothic letter \(a\), and its elements by \(a(nm)\). That is, they express the matrix \(a\) in the following way:

\[
a = (a(nm)) = \begin{pmatrix}
a(00) & a(01) & a(02) & \cdots \\
a(10) & a(11) & a(12) & \cdots \\
a(20) & a(21) & a(22) & \cdots \\
\vdots & \vdots & \vdots & \ddots 
\end{pmatrix}
\]

\(a\) is taken to be a square infinite matrix.

\(^{30}\)M. Born und P. Jordan, Zeit. f. Phys. 34 (1925), 858.

\(^{31}\)M. Bôcher, "Einführung in die höhere Algebra" (Teubner, Leipzig, 1910; Translation by H. Beck from the English original).

Born-Jordan first treat a system having a single degree of freedom, as in the case of Heisenberg's paper. They start the second chapter of their paper "Dynamics", by saying that "a dynamical system is to be described by the coordinate $q$ and momentum $p$", and "these should be represented by the matrices

$$q = (q(nm)e^{2\pi i\nu(nm)t}), \quad p = (p(nm)e^{2\pi i\nu(nm)t}),$$

(40)
in which $\nu(nm)$ means the quantum-theoretical frequency which belongs to the transition between the states with the quantum numbers $n$ and $m$. In this way, Born-Jordan presuppose, like Heisenberg, the existence of the quantum number, basing on the results of the old quantum mechanics. This is in a sharp contrast to the case of Schrödinger's wave mechanics, the aim of which is to construct a theory independent of the presupposition of existence of quantum numbers, as we shall see in the next chapter.

The matrices given in Eq. (40) "should be Hermitian", and this condition "should hold for all real $t$". One has therefore

$$a(nm)q(mn) = |q(nm)|^2, \quad (41)$$

$$\nu(nm) = -\nu(mn). \quad (42)$$

"If $q$ is a Cartesian coordinate", $|q(nm)|^2$ is "a measure of the probability of the transition $n \rightarrow m". This corresponds to the assumption in Heisenberg's theory that the quantum-theoretical amplitudes of $-ex(t)$ are to give the transition probabilities.

Born-Jordan "want further to demand" that

$$\nu(jk) + \nu(kl) + \nu(lj) = 0. \quad (43)$$

From this equation, together with Eq. (42), it follows that there exist such quantities $W_n$ that

$$h\nu(nm) = W_n - W_m. \quad (44)$$

These two equations of Born-Jordan represent Ritz's combination rule and Bohr's frequency relation, respectively. Born-Jordan point out that, on account

33) As is well known, from the Hermitian property of $q$ it results that $q(nm)e^{2\pi i\nu(nm)t} = q(mn)e^{-2\pi i\nu(mn)t}$. Because this relation holds for any real $t$, Eqs. (41) and (42) are derived.
of Eq. (44), a function \( g(p, q) \) which can be developed in a power series of \( p \) and \( q \) takes also the form

\[
g = (g(nm)e^{2\pi i\nu(nm)t}).
\]

(45)

This is nothing but the representation in matrix form of the combination rule of frequencies, and of Heisenberg's combination rule of quantum-theoretical amplitudes given in Eq. (12) in the preceding section. Because of Eq. (45), Born-Jordan adopt the abbreviated notation

\[
q = (q(nm)), \quad p = (p(nm)),
\]

(46)

instead of Eq. (40).

The time derivative \( \dot{g} \) of \( g \) is obtained from Eq. (45) to be

\[
\dot{g} = 2\pi i(\nu(nm)g(nm)),
\]

(47)

in abbreviated notation. If \( \nu(nm) \neq 0 \) for \( n \neq m \), \( \dot{g} = 0 \) means that \( g \) is a diagonal matrix.

Born-Jordan assume that if the Hamiltonian is of the form

\[
H = \frac{p^2}{2m} + U(q),
\]

(48)

there hold the equations of motion of the same form as is given in the classical theory

\[
\hat{q} = \frac{\partial H}{\partial p} = \frac{p}{m}, \quad \hat{p} = -\frac{\partial H}{\partial q} = -\frac{\partial U}{\partial q}.
\]

(49)

It is necessary, however, to take an arbitrary \( H(p, q) \), "in consideration of the relativistic mechanics and particularly of the treatment of the motion of electrons under the action of magnetic fields". Born-Jordan make a "study in accordance with the correspondence principle" of the action principle in the classical theory

\[
\int_{t_0}^{t_1} Ldt = \int_{t_0}^{t_1} \{p\dot{q} - H(p, q)\}dt = \text{Extremum}.
\]

(50)

\text{34)} The derivation of a function of matrixes considered here, is to be carried out with respect to each matrix term in the power series development of the function, after putting each term in its full form of multiplication. For example, if \( y = x_2^2x_2x_1x_3 \), \( \partial y/\partial x_1 = \partial x_2^2x_2x_1x_3/\partial x_1 = x_2^2x_1x_3 + x_2^2x_1x_3 + x_3x_1x_2 \).
If $L$ in Eq. (50) is expanded as a Fourier series and the time interval $t_1 - t_0$ is taken sufficiently large, only the constant term of $L$ contributes to the action integral. "The form, which the action principle thus obtains, suggests the following translation into quantum mechanics". That is, the diagonal sum of $L$, $D(L) = \sum_k L(kk)$, should be brought to be an extremum. To be precise,

$$ D(L) = D(p\dot{q} - H(p, q)) = \text{Extremum}, \tag{51} $$

for suitably chosen $p$ and $q$ and fixed $\nu(nm)$. From Eq. (51) it is deduced that

$$ \ddot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}. \tag{52} $$

One sees that the equation of motion is given in the canonical form in quantum mechanics also.

§2. Presentation of the essentially complex commutation relation

Born-Jordan write "the quantum condition in the 'classical' quantum theory" as

$$ J = \int pdq = \int_0^{1/\nu} p\dot{q}dt, \tag{53} $$

and, with the use of the Fourier developments

$$ p = \sum_{\tau = -\infty}^{+\infty} p_\tau e^{2\pi i \nu \tau}, \quad q = \sum_{\tau = -\infty}^{+\infty} q_\tau e^{2\pi i \nu \tau}, \tag{54} $$

put it in the form by the differentiation with respect to $J$

$$ 1 = 2\pi i \sum_{\tau = -\infty}^{+\infty} \tau \frac{\partial}{\partial J} (q_\tau p_{-\tau}). \tag{55} $$

If $p = m\dot{q}$, this equation becomes equivalent to Eq. (17) in the preceding section, as $p_{-\tau}$ can be given in terms of $q_{-\tau}$. As we have seen in the preceding section, Eq. (17) is transformed by Heisenberg into his quantum condition.

However, Born-Jordan say that "as the assumption $p = m\dot{q}$ should not be made here, we must translate Eq. (55) directly into a difference equation".\footnote{The original number of Eq. (55) in Born-Jordan's paper is (36).}
The summation in Eq. (55) "should correspond to"

\[
\frac{1}{\hbar} \sum_{\tau = -\infty}^{+\infty} \{ q(n + \tau, n)p(n, n + \tau) - q(n, n - \tau)p(n - \tau, n) \}
\]

in which every \(q(nn)\) or \(p(nm)\) is to be put to zero, if it has a negative index. They get thus "the quantum condition"

\[
\sum_k \{ p(nk)q(kn) - q(nk)p(kn) \} = \frac{\hbar}{2\pi i}, \tag{56}
\]

"as the corresponding revision" of Eq. (55). This quantum condition becomes

\[
\sum_k \nu(kn)|q(nk)|^2 = \hbar/8\pi^2m \text{ if } p = m\dot{q},
\]

coinciding with Heisenberg's quantum condition, Eq. (18), in the preceding section.\(^{36}\)

Now, for the matrix \(d = pq - qp\), it can be shown by making use of Eq. (52) that \(d = 0\). \(d\) is therefore a diagonal matrix, whose diagonal elements are given by Eq. (56). One has thus

\[
pq - qp = \left( \frac{\hbar}{2\pi i} \right) 1, \tag{57}
\]

where 1 denotes the unit matrix. Born-Jordan call this relation the "intensified quantum condition" (verschärfte Quantenbedingung). This is a fundamental relation that is later called the commutation relation between the momentum and the coordinate.

Heisenberg's "quantum-theoretical quantum condition" now gives the diagonal elements of this relation in the particular case of \(p = m\dot{q}\). Heisenberg's "quantum-theoretical quantum condition" is generalized by Born-Jordan to their "intensified quantum condition" through consideration of the general case in which \(p\) is not given by \(m\dot{q}\) and by their application of the variational principle to the matrix \(L\) corresponding to the Lagrangian to get the equation of motion in matrix form.

In Heisenberg's work, the quantum-theoretical amplitudes of a physical quantity corresponding to various transitions are taken to be real, as is seen from Eq. (5) in the preceding section,

\[
R\{ A(n, n - \alpha)e^{i\omega(n, n - \alpha)t} \}.
\]

In his search for the combination rule of amplitudes, however, "the aggregate of amplitudes" is taken into consideration, and a multiplication rule is introduced for complex amplitudes \(A(n, n - \alpha)e^{i\omega(n, n - \alpha)}\) themselves, instead of their real

\(^{36}\)Note that \(\nu(kn) < 0\) for \(k < n\).
parts, on the supposition that "the phase also has a definite meaning in quantum mechanics". The combination rule of quantum-theoretical amplitudes, Eq. (10) in the preceding section, which is in accordance with Ritz's combination rule, could not be arrived at if \( \cos\{\omega(n, n - \alpha)t\} \) or \( \sin\{\omega(n, n - \alpha)t\} \) were used. Heisenberg's "aggregate of quantum-theoretical amplitudes" is grasped in Born-Jordan's work as a matrix having these amplitudes as its elements.

In the classical theory, Fourier series developments into complex exponential functions, such as Eq. (13) or (54), are employed for the sake of convenience, and the relations \( a_{-\alpha}(n) = \tilde{a}_\alpha(n) \), \( p_{-\tau} = \tilde{p}_\tau \), etc. are to represent the real property of the physical quantities concerned. Born-Jordan translate these relations into the Hermitian property of matrices and, at the same time, translate the quantum condition in the old quantum mechanics, Eq. (55), which is expressed with the use of \( i \) for convenience of calculation, into the commutation relation, Eq. (57), which involves \( i \) essentially. In this way, the important lawfulness of Ritz's combination rule, Bohr's frequency relation and Sommerfeld's quantum condition reached in the old quantum mechanics is presented anew, being raised to the commutation rule between the coordinate and momentum represented by matrices, through the mediation of Heisenberg's combination rule of quantum-theoretical amplitudes.\(^{37}\)

\[ §3. \textit{The laws of time variation and of energy conservation in matrix mechanics} \]

As we have seen in §4 of the preceding section, Heisenberg applied his new quantum condition to a nonharmonic oscillator to get an approximate solution, and showed that the "oscillating terms" of the energy \( W \) was zero, and the "constant term" satisfied Bohr's frequency relation. But, he was not able to prove this generally. In Born-Jordan's formulation, this corresponds to the diagonality of the matrix \( H \) of the Hamiltonian satisfying the equation \( \dot{H} = 0 \), which is proved using their "intensified quantum condition".

It results from Eq. (57) by means of the inductive method that

\[ p^n q = qp^n + \left( \frac{\hbar}{2\pi i} \right) np^{n-1}. \]  

\(^{37}\)On this point, cf. also the end of Sec. 4, §3, Chap. 2 of the present volume.
Because Eq. (57) is invariant for the interchange of $p$ and $h$ with $q$ and $-h$, respectively, it is also deduced that

$$q^n p = pq^n - \left( \frac{h}{2\pi i} \right) nq^{n-1}. \quad (58')$$

If the Hamiltonian is given by

$$H = H_1(p) + H_2(q) \quad (59)$$

and each of $H_1$ and $H_2$ can be developed into a power series of $p$ and $q$, Eqs. (58) and (58') give

$$Hq - qH = \left( \frac{h}{2\pi i} \right) \frac{\partial H}{\partial p} \quad (58')$$
$$Hp - pH = - \left( \frac{h}{2\pi i} \right) \frac{\partial H}{\partial q}. \quad (60)$$

Inserting these equations into the canonical equation of motion, one has

$$\dot{q} = \left( \frac{2\pi i}{h} \right) (Hq - qH), \quad \dot{p} = \left( \frac{2\pi i}{h} \right) (Hp - pH). \quad (61)$$

Next, if the notation

$$\begin{vmatrix} H \\ g \end{vmatrix} = Hg - gH \quad (62)$$

is used, one has the relation

$$\begin{vmatrix} H \\ ab \end{vmatrix} = \begin{vmatrix} H \\ a \end{vmatrix} b + \begin{vmatrix} H \\ b \end{vmatrix} a \quad (63)$$

It is concluded, therefore, by making use of the power-series expansion of a matrix function $g(p, q)$ in $p$ and $q$, that

$$\dot{g} = \frac{2\pi i}{h} \begin{vmatrix} H \\ g \end{vmatrix} = \frac{2\pi i}{h} (Hg - gH). \quad (64)$$

One then gets $\dot{H} = 0$ by putting $g = H$. $H$ turns out therefore to be diagonal. From Eqs. (47) and (61), it is deduced that

$$h\nu(nm)q(nm) = \{H(nn) - H(mm)\}q(nm), \quad (65)$$

which gives Bohr's frequency relation.
For a Hamiltonian more general than that given in Eq. (59), Born-Jordan use $H^* = p^2 q$, for example, for which $\dot{H}^* \neq 0$. $H^*$ gives, however, the same equation of motion as $H = (p^2 q + q p^2)/2$ for which $\dot{H} = 0$. It can be shown that this also holds for

$$H^* = p^s q^r, \quad H = \sum_{l=0}^{s} p^{s-l} q^r p^l/(s+1), \quad (66)$$

with the use of Eqs. (58) and (58'). The $H$ given by Eq. (66) satisfies Eq. (60), so that Eq. (64) holds for this $H$. For such an arbitrary $H^*(p, q)$ which is given by a power series in $p$ and $q$, the same thing can be said on account of the linearity in $H^*$ and $H$ of the conditions $\partial H^*/\partial p = \partial H/\partial p$ and $\partial H^*/\partial q = \partial H/\partial q$, to be satisfied for the equivalence of $H^*$ and $H$ in giving the same equation of motion.

In this way, Born-Jordan prove the validity of the conservation law of energy and Bohr's frequency relation in matrix mechanics. Equation (64), proved in the course of their proof becomes, when expressed using the symbol $[,]$ for the commutation relation and $h = h/2\pi$, as in present-day quantum mechanics, nothing but

$$\dot{g} = (i/h)[H, g] = (i/h)(Hg - gH),$$

which is called the Heisenberg equation of motion and constitutes one of the most fundamental laws in quantum mechanics.\(^{38}\)

Detailed analyses are given by Born-Jordan, on the basis of their formulation using matrices of Heisenberg's theory, for the examples of nonharmonic and harmonic oscillators.

§4. Generalization by Born-Heisenberg-Jordan to systems of multiple degrees of freedom

Born, Heisenberg and Jordan generalize, in their paper received on Nov. 16, 1925 “On Quantum Mechanics II”,\(^{39}\) Born-Jordan's formulation to systems of multiple degrees of freedom. They also develop the perturbation theory in

\(^{38}\)As will be described in the next subsection, Dirac arrived at this equation independently of Born and Jordan. This equation should be more appropriately called the Born-Dirac-Jordan equation of motion. It is called the Heisenberg equation of motion, perhaps because of its contrast to the Schrödinger equation of motion arrived at in wave mechanics.

quantum mechanics with the use of the "canonical transformation", prove the conservation law of linear and angular momenta, and treat the selection rule and relative intensities in the Zeeman effect.

Born-Heisenberg-Jordan say that "for \( f > 1 \) degrees of freedom it suggests itself to choose, instead of a representation of quantum-theoretical quantities by two-dimensional matrices, one by \( 2^f \)-dimensional matrices, corresponding to the \( 2^f \)-dimensional manifold of the state in the classical \( J \)-space:

\[
\begin{align*}
q_k &= (q_k(n_1 \cdots n_f, m_1 \cdots m_f)), \\
p_k &= (p_k(n_1 \cdots n_f, m_1 \cdots m_f))
\end{align*}
\tag{67}
\]

where \( k = 1, 2, \ldots, f \). Their words are based on the result reached in the old quantum mechanics that every stationary state of a system of \( f \) degrees of freedom is specified by \( f \) quantum numbers.

They point out that "this representation is however by no means essential". This is because "one can carry out any arbitrary rearrangement, and in particular convert the \( 2^f \)-dimensional matrices also into two-dimensional matrices". "The basic definitions of addition and multiplication, as well as differentiation with respect to time, are clearly quite independent of any ordering relations between the systems of indexes \( n_1, \ldots, n_f \), which are taken to indicate singly the states and in pairs the transitions". That is, it suffices to number the systems by indexes \( n_1, \ldots, n_f \).

Accordingly, the variational principle introduced by Born-Jordan for the diagonal sum of the matrix of Lagrangian \( L \) can be applied to the problem under consideration, so that one has the set of equations

\[
\begin{align*}
\dot{q}_k &= \frac{\partial H}{\partial p_k}, \\
\dot{p}_k &= -\frac{\partial H}{\partial q_k}.
\end{align*}
\tag{68}
\]

Born-Heisenberg-Jordan say that, as an "obvious generalization" of the commutation relation ("Vertaushungsrelation"), there "come" the equations

\[
\begin{align*}
p_k q_l - q_l p_k &= \left( \frac{\hbar}{2\pi i} \right) \delta_{kl}, \\
p_k p_l - p_l p_k &= 0, \\
q_k q_l - q_l q_k &= 0.
\end{align*}
\tag{69}
\]

By means of these commutation relations, the conservation law of energy can be derived, as in the case of a system of a single degree of freedom.

Born-Heisenberg-Jordan consider a transformation of the set of variables \( p_k, q_k \) into another set of variables \( P_k, Q_k \), and call it the "canonical
transformation" when the new $P_k, Q_k$ satisfy commutation relations of the same kind as Eq. (69). This is because, if it is so, the equations of motion of the same kind are valid for $P_k, Q_k$, just as those for $p_k, q_k$. They show that with an "orthogonal matrix" $S$ satisfying the relations

\[ SS^* = 1, \quad S^* S = 1, \quad (70) \]

such a canonical transformation is given by

\[ P_k = S p_k S^*, \quad Q_k = S q_k S^*. \quad (71) \]

By this transformation, the Hamiltonian $H(p_k, q_k)$ is also transformed into

\[ H(P_k, Q_k) = SH(p_k, q_k)S^*. \quad (72) \]

If $H(P_k, Q_k)$ becomes a diagonal matrix $W$, the energy of the system is known from the diagonal elements of $W$. Born-Heisenberg-Jordan point out that the problem of finding such an $S$ is equivalent to that of finding the transformation into the principal axes of the "Hermitian form"

\[ A(x, x^*) = \sum_{mn} H(mn) x_m x_n^*, \quad (73) \]

that is, to the problem of finding the solution of the homogeneous simultaneous equations

\[ W x_m - \sum_n H(mn) x_n = 0 \quad (m = 1, 2, 3, \ldots). \quad (74) \]

Equation (74) is mathematically identical with the equation that is called 'eigenvalue equation of energy in matrix form' in present-day quantum mechanics. It is the wave mechanics, however, that brings on the key to grasp the physical meaning of this equation, as we shall see in the next chapter.

Now, if $H(mn)$ is given by a power series in a parameter $\lambda$ as

\[ H(mn) = W^0 \delta_{mn} + \lambda H_1(mn) + \lambda^2 H_2(mn) + \cdots, \quad (75) \]

\[ ^{40} \tilde{S} \text{ and } S^* \text{ mean, respectively, the transposed and complex conjugate of } S, \text{ so that } S^* \text{ is identical to the Hermitian conjugate } S^d \text{ of } S. \text{ In present day textbooks on quantum mechanics, Eq. (70) is called the unitarity of } S, \text{ and Eq. (71) a unitary transformation.} \]

\[ ^{41} H(p_k, q_k) \text{ is the abbreviation of } H(p_1, q_1, \ldots, p_k, q_k, \ldots), \text{ and similarly is } H(P_k, Q_k). \]

\[ ^{42} \text{When the coordinates } x_1, x_2, \ldots \text{ are transformed into the coordinates along the principal axes, } A(x, x^*) \text{ takes the form } A(x', x'^*) = \sum_n W_n x'_n x'^*_n. \text{ If the } n\text{th solution of Eq. (74) which is normalized to unity is denoted as } x_{mn}, \text{ one has } S = (x_{mn}). \]
Eq. (74) can be solved in a recursive way to give \( W \) and \( x_m \) in power series form:

\[
\begin{align*}
W &= W^{(0)} + \lambda W^{(1)} + \lambda^2 W^{(2)} + \cdots , \\
x_m &= x_m^{(0)} + \lambda x_m^{(1)} + \lambda^2 x_m^{(2)} + \cdots .
\end{align*}
\] (76)

This is the "perturbation method" given by Born-Heisenberg-Jordan, which has the identical form as the perturbation formula in present-day quantum mechanics.

§5. Angular momenta and the Zeeman effect in matrix mechanics

From the commutation relations, it turns out that, with a Hamiltonian of the form \( H(p_k, q_k) = H(p_k) + H(q_k) \), the \( x \)-, \( y \)- and \( z \)-components of the momentum of the system, \( p_x, p_y \) and \( p_z \), and those of the angular momentum of the system, \( M_x, M_y \) and \( M_z \), have zero derivatives with respect to time. Also from the commutation relations, it can be deduced that

\[ M_x M_y - M_y M_x = i \left( \frac{\hbar}{2\pi} \right) M_z . \] (77)

Born-Heisenberg-Jordan say that "the area theorem, as in the classical theory, holds for at most one or for all the three axes".

Even when an atom is acted on by a force, one has \( \dot{M}_z = 0 \) if the force is symmetrical about the \( z \)-axis. In this case, because \( M_z \) is diagonal and \( q_{lz} M_z - M_z q_{lz} = 0 \) if the \( z \)-coordinate of the \( l \)th electron is \( q_{lz} \), one gets

\[ q_{lz}(mn)(M_z n - M_z m) = 0. \] (78)

Consequently, "for a quantum jump in which the angular moment \( M_z \) changes, the 'plane of oscillation' of the generated 'spherical wave' lies perpendicular to the \( z \)-axis". From the commutation relation between \( M_z \) and \( q_{lx} \) or \( q_{ly} \), where \( q_{lx} \) and \( q_{ly} \) are the \( y \)- and \( z \)-coordinates respectively of the \( l \)th electron, one gets

\[ ((M_z n - M_z m)^2 - (h/2\pi)^2)q_{\eta}(mn) = 0; \quad \eta = x \text{ or } y . \] (79)

Therefore, "for a quantum jump without any change in \( M_z \) the emitted light is linearly polarized parallel to the \( z \)-axis". Thus, the rules for the polarization
of emitted light, which were derived by means of the correspondence principle in the old quantum mechanics, are confirmed by matrix mechanics.

Since $M_z$ changes by $\pm h/2\pi$ according to Eq. (79), one has

$$M_{zn} = (h/2\pi)(n_1 + C); \quad n_1 = 0, \pm 1, \pm 2, \ldots,$$

(80)
in terms of the particular quantum number $n_1$ among the set of the quantum numbers, that is involved in $M_z$. Thus it results from the diagonal sum of the z-component of Eq. (77) that

$$\sum_{n_1} \left( \frac{h}{2\pi} \right) (n_1 + C) = 0.$$

(81)

$n_1 + C$ should lie symmetrically with respect to zero, so that $m = n_1 + C$ is to be an integer or a half-integer and obeys the selection rule $m \rightarrow m, m \pm 1$. This supports, on the basis of fundamental laws, Landé’s introduction in 1921, described in Sec. 1, §1, Chap. 4, Vol. II, of half-integer values besides integer values of $m$, from consideration of the symmetric positive and negative values of $m$, which should change by unity according to the selection rule.

Born-Heisenberg-Jordan show, furthermore, that the diagonal elements of $M^2$ are given by $M^2 = j(j+1)(h/2\pi)^2$ with $j = m_{\text{max}}$, so that $j$ is an integer or a half-integer and obeys the selection rule $j \rightarrow j, j \pm 1$, by making use of the commutation relations between $M^2, M_z, Z_i = q_{lx}M_x + q_{ly}M_y + q_{lz}M_z$ and $X_i = q_{ly}M_x - M_yq_{lx}$, the properties of $M_x \pm iM_y$, and others. This justifies, also from fundamental laws, Sommerfeld’s expression $j(j+1)$, in units of $(h/2\pi)^2$, for the squared value of the angular momentum of an electron, that was due to Landé’s corresponding expression $J^2 = (J - 1/2)(J + 1/2)$ introduced in connection with the problem of $g$-factor.

Born-Heisenberg-Jordan also calculate, for the cases of $j \rightarrow j \pm 1$, the ratios among the matrix elements of $q_{lx}, q_{lx} + iq_{ly}$ and $q_{lx} - iq_{ly}$, which obey in each case the selection rules $m \rightarrow m, m \rightarrow m + 1$ and $m \rightarrow m - 1$, respectively. These ratios are to give “the intensity formulae for the Zeeman effect”. This is however about the normal Zeeman effect. They say that “we must

---

43) $Z_i$ corresponds classically to the scalar product between the position vector of the $i$th electron and the angular momentum of the system, and $X_i$ the $z$-component of the vector product between them.

44) The procedure of analyses used is almost the same as that given in present-day textbooks on quantum mechanics.

45) Cf. Sec. 1, §2, (ii), Chap. 4, Vol. II.
emphasize that a series of more weighty reasons speak against the possibility of any interpretation of the anomalous Zeeman effect on this basis”. "Rather, one may perhaps be allowed to hope that Uhlenbeck-Goudsmit's assumption later permits a quantitative description of the mentioned phenomena”.

Saying so, Born-Heisenberg-Jordan cite Uhlenbeck-Goudsmit's paper\(^{46}\) published about a month earlier, in which the "eigen-rotation" of the electron was introduced in relation to the anomalous coefficient of the magnetic moment of certain atoms. As have been mentioned in Sec. 1, §1 of the present chapter, Heisenberg pursued, basing on the results of the old quantum mechanics in which the various contradictions were investigated, systematized and clarified through intricate processes, “a quantum-theoretical mechanics analogous to the classical mechanics”, in which “relations between observable quantities appear only”. Heisenberg's idea is formulated by Born-Jordan into matrix mechanics, which is further developed by Born-Heisenberg-Jordan to deduce theoretically, the necessity of half-integer, besides integer, quantum numbers, for angular momentum. However, for the anomalous Zeeman effect itself, from which originated the introduction of half-integer quantum numbers for angular momentum, it does not suffice to consider “relations between observable quantities” only, and so Born-Heisenberg-Jordan have to assume the substantialistic model proposed by Uhlenbeck-Goudsmit. That is, for the study of this phenomenon they necessitate a new clue for the substantial structure of the object.

Heisenberg and Jordan investigate, in their paper received on Mar. 16, 1926 “Application of Quantum Mechanics to the Problem of the Anomalous Zeeman Effect”,\(^{47}\) “the quantum mechanical behavior of the atomic model characterized by the hypothesis of the rotating electron”. Denoting the Hamiltonian of a hydrogen-like atom by \(H_0\), the angular momentum of the valence electron by \(\vec{k}\), the angular momentum of its eigen-rotation by \(\vec{s}\), its position relative to the nucleus by \(\vec{r}\), and the electric charge of the nucleus by \(Ze\), they write the Hamiltonian of the atom in an external magnetic field \(\vec{H}\) as\(^{48}\)

\[
\begin{align*}
H &= H_0 + H_1 + H_2, \\
H_1 &= \left(\frac{e}{2mc}\right) \vec{H} \cdot (\vec{k} + \vec{2}), \quad H_2 = \left(\frac{Ze^2}{2m^2c^2}\right) (1/r^3) \vec{k} \cdot \vec{s}.
\end{align*}
\]

\(^{46}\)G. E. Uhlenbeck und S. Goudsmit, Die Naturwiss. 13 (1925), 953 [the same as Ref. 72) in Sec. 3, §3, Chap. 4, Vol. II.


\(^{48}\)“means a vector. \(H, \vec{k}, \vec{s}\), etc. are matrices, which would be expressed as \(\vec{H}, \vec{k}, \vec{s}\), etc. in the notation of Born et al.
The coefficient 2 in the right-hand side of $H_1$ is the anomalous coefficient of the magnetic moment of the electron. $H_2$ is the interaction energy between the magnetic field generated by the motion of the electron relative to the nucleus and the anomalous magnetic moment of the electron, for whose strength use is made of the value due to Thomas.\(^{49}\)

$H_0$ is independent of the direction of $\vec{s}$. Heisenberg-Jordan denote the $z$-components of $\vec{k}$, $\vec{s}$ and $\vec{M} = \vec{k} + \vec{s}$ respectively by $k_z = m_k (h/2\pi)$, $s_z = m_s (h/2\pi)$ and $M_z = m (h/2\pi)$, where $m = m_k + m_s$, and decompose Eq. (74), for the problem under consideration, into sets of two-dimensional simultaneous equations, each set corresponding to a specified value of $m$ and $m_s = \pm 1/2$. They find from each set the “canonical transformation” to diagonalize $H$. The two values of $W$ obtained from each set are the “diagonal terms” of $H$, which give the splitting of the spectral term of the respective value of $m$ in the anomalous Zeeman effect. They obtain also the ratios among the intensities of the lines in the Zeeman effect by calculating the matrix elements of $Q_z$ and $Q_x \pm iQ_y$ from those of $q_z$ and $q_x \pm iq_y$ given by Born-Heisenberg-Jordan, with the use of the canonical translator of the form of Eq. (71).\(^{50}\)

In the case of a weak magnetic field, if $H_1$ is treated as a perturbation, the magnetic moment makes a precession around the total angular momentum $\vec{M} = \vec{k} + \vec{s}$, so that the energy due to the magnetic field is given by

$$\tilde{H}_1 = \left( \frac{e}{2mc} \right) \mathcal{H} \cdot \vec{M} (\vec{k} + 0\vec{s}) \cdot \vec{M} \vec{M}^0$$

\[ \text{(83)} \]

$$= \left( \frac{e}{2mc} \right) \mathcal{H} m \left\{ 1 + \frac{\vec{M} \cdot \vec{s}}{\vec{M}^2} \right\}$$

\[ \text{(84)} \]in consideration of the projection of the magnetic moment in the direction of $\vec{M}$. Heisenberg-Jordan point out that, with the use of the formulae $\vec{M}^2 = (h/2\pi)^2 j(j + 1)$, $\vec{k}^2 = (h/2\pi)^2 k(k + 1)$ and $\vec{s}^2 = (h/2\pi)^2 s(s + 1)$ shown by them together with Born, the quantity in $\{ \}$ of Eq. (84) gives Landé’s $g$-factor.

As we have described in Sec. 1, §1 and §2, Chap. 4, Vol. II, Landé systematized in 1921 the results of experiments on the anomalous Zeeman effect by taking half-integer values for Sommerfeld’s inner quantum number, and Heisenberg attempted in 1922 to explain the half-integer quantum numbers by

\(^{49}\) \(\langle \rangle\) means the average value [cf. Sec. 3, §3, Chap. 4, Vol. II]. $H_2$ is what is called the spin-orbit interaction energy nowadays.

\(^{50}\) Their procedure of calculation is almost the same as those seen in present-day textbooks on quantum mechanics. In Heisenberg-Jordan’s calculation of $1/r^3$, however, matrix mechanics is applied on the assumption that the electron makes a motion in a given plane.
means of a substantialistic model, such that the angular momentum of the electron was divided to the atomic body (atomic core) on account of their mutual interaction. After the subsequent process as has been described in Secs. 2 and 3, Chap. 4, Vol. II, Uhlenbek-Goudsmit’s substantialistic model was presented in which the half-integer quantum number, attributed so far to the atomic body, was considered, rather, to result from the eigen-rotation of the electron. It is interesting to note that Heisenberg, who intended to treat the relations among observable quantities, applied, in collaboration with Jordan, matrix mechanics developed by Born and Jordan in collaboration with Heisenberg to the substantialistic model of Uhlenbeck-Goudsmit.

§6. Matrix mechanics and the oscillation of a string — the beginning of quantum field theory

Born and Jordan attempt to treat electrodynamics with matrix mechanics, in their first paper “On Quantum Mechanics”\(^{51}\) mentioned in §1 of the present section. Saying that “one will represent electromagnetic phenomena as a superposition of plane waves”, Born-Jordan regard the electric and magnetic fields as matrices. For example, they consider the electric field to be represented by\(^{52}\)

\[
\vec{E} = (\vec{E}(nm)e^{2\pi i\nu(nm)(t-x/c)}).
\]

The following Maxwell equations are also considered to be valid in matrix form:

\[
\text{rot} \vec{H} - \frac{\dot{\vec{E}}}{c} = 0, \quad \text{rot} \vec{E} - \frac{\dot{\vec{H}}}{c} = 0.
\]

With the definition of the energy density \(W = (\vec{E}^2 + \vec{H}^2)/8\pi\), they get from Eq. (86) by differentiation of matrices\(^{53}\)

\[
\dot{W} + \text{div} \vec{S} = 0; \quad \vec{S} = (c/8\pi)(\vec{E} \times \vec{H} - \vec{H} \times \vec{E}).
\]

Born-Jordan say that “this is the Poynting law for matrix electrodynamics”. They introduce the matrixized Hertz vector \(\vec{Z}\), which should satisfy

\(^{51}\)M. Born und P. Jordan, Zeit. f. Phys. 34 (1925), 858 [the same as Ref. 28] in the present chapter.

\(^{52}\)Here a Gothic letter denotes a matrix and the symbol \(\sim\) a vector.

\(^{53}\)In the classical electrodynamics there holds the relation \(\vec{E} \times \vec{H} = -\vec{E} \times \vec{H}\), but after being matrixized it becomes \(\vec{E} \times \vec{H} \neq -\vec{H} \times \vec{E}\) in general.
the relations as in the classical theory,

\[
\vec{E} = \text{grad} \ \text{div} \ \vec{Z} - \frac{\vec{Z}}{c^2}, \quad \vec{H} = \text{rot} \ \frac{\vec{Z}}{c}.
\]  

(88)

In the classical theory, for a spherical wave \( \vec{Z} \) is proportional to \((1/r)e^{2\pi i\nu(t-r/c)}\). From this fact, Born-Jordan assumes that \( \vec{Z} \) is given in terms of the electric dipole moment \( e\vec{q} \) by

\[
\vec{Z} = (e\vec{q}(nm)(1/r)e^{2\pi i\nu(nm)(t-r/c)}).
\]  

(89)

Inserting this into Eq. (88) and then into Eq. (87), they get for the energy flowing out of a region per unit time the expression

\[
\int \vec{S} \cdot d\vec{\sigma} = \left(\frac{2e^2}{3c^2}\right) \vec{q}^2,
\]  

(90)

where \( d\vec{\sigma} \) denotes the surface element vector of the region.

They replace \( \vec{q} \) with \( q \) in consideration of an oscillation of the dipole moment in one direction, and then take the diagonal sum of Eq. (90), instead of the time average in classical theory, to derive

\[
D \left\{ \left(\frac{2e^2}{3c^2}\right) \vec{q}^2 \right\} = \left(\frac{64\pi^4e^2}{3c^3}\right) \sum_n \sum_{k<n} \nu(nk)^4 |q(nk)|^2.
\]  

(91)

As the right-hand side of Eq. (91) is given as a sum over the states indicated by \( k < n \) for each state indicated by \( n \), Born-Jordan conclude that Heisenberg's assumption of the quantum-theoretical amplitude \( q(nk) \) as being the determinant of the transition probability "is in a restricted meaning justified".

Equation (89) used by Born-Jordan is not the one that is derived from the solution of the matrixized Maxwell equations for the case of a charged particle as the source, but is what is assumed in accordance with the correspondence principle from the solution of the classical Maxwell equations in that case. Notice is to be taken, however, of Born-Jordan's attempt itself of formulating electrodynamics in the form of matrix mechanics.

Born-Heisenberg-Jordan apply matrix mechanics, on the other hand, to the calculation of the fluctuation in energy of an oscillating string, in their paper "On Quantum Mechanics II" cited above.\(^{54}\) In classical theory, the oscillation

\(^{54}\)M. Born, W. Hwisenberg und P. Jordan, Zeit. f. Phys. \textbf{35} (1926), 557 [the same as Ref. 38] in the present section.
The Proposal and Formulation of Matrix Mechanics

of an elastic body such as crystalline lattice, for example, can be treated by means of the eigen oscillations of the atoms at the lattice points. It can also be treated by idealizing the oscillation of the elastic body as being continuous in consideration of its extremely large degree of freedom. In this case, the fluctuation in energy results from "the interference of wave fields". As was shown by Einstein, however, the interference of wave fields in the case of cavity radiation gives the fluctuation in energy that is characteristic of the Rayleigh-Jeans formula of radiation.\textsuperscript{55) Thus, Born et al. "calculate the fluctuation in energy from the interference in accordance with quantum mechanics".}

Born-Heisenberg-Jordan "take the conceivable simplest case, namely a stretched oscillating string, in order to avoid complexities in calculation, which do not affect the essence of the matter". The string is supposed to be stretched between $x = 0$ and $x = l$. At first, they develop the displacement $u(x, t)$ of the string in Fourier series as

$$u(x, t) = \sum_{k=1}^{\infty} q_k(t) \sin \omega_k x, \quad \omega_k = \frac{k\pi}{l}, \quad (92)$$

according to the classical theory. The energy of the oscillatory string, $H$, can be written as\textsuperscript{56)\textsuperscript{56)\textsuperscript{56})}

$$H = \frac{1}{2} \int_{0}^{l} \left\{ \dot{u}^2 + \left( \frac{\partial u}{\partial x} \right)^2 \right\} dx = \frac{l}{4} \sum_{k=1}^{\infty} (\dot{q}_k^2 + \omega_k^2 q_k^2). \quad (93)$$

The energy $E$ contained in the part of the string between $x = 0$ and $a$ is obtained by replacing the upper limit of the integration in Eq. (93) from $l$ to $a$, so that it is expressed as

$$E = \frac{l}{4} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (\dot{q}_j \dot{q}_k K_{jk} + \omega_j \omega_k q_j q_k K'_{jk}), \quad (94)$$

where

$$\begin{cases} K_{jk} = \frac{\sin(\omega_j - \omega_k) a}{\omega_j - \omega_k} \mp \frac{\sin(\omega_j + \omega_k) a}{\omega_j + \omega_k} \\ K'_{jk} \end{cases} \quad (95)$$

$q_k$ oscillates as $q_k = A_k \cos(\omega_k t + \varphi_k)$. If the average value of a quantity over the phase $\varphi_k$ is denoted by the symbol $\bar{\cdot}$, one has $\bar{E} = (a/l)H$. The

\textsuperscript{55)} Cf. Sec. 3, §1, Chap. 1, Vol. II.

\textsuperscript{56)} Born et al. set the density and the tension to be unity.
fluctuation of $E$, that is $\Delta = E - \bar{E}$, is equal to the double summation given in Eq. (94) taken over only terms with $j \neq k$. If $l$ is very large and $a$ is also large enough, the summation over $k$ in the expression for $\Delta$ can be replaced with integration with respect to $\omega_k$. One gets then

$$\bar{\Delta}^2 = \left( \frac{al^2}{8\pi} \right) \left[ \int_0^{\infty} d\omega \left\{ \frac{(\bar{q}^2_\omega)^2}{2} + (\omega^2 q^2_\omega)^2 + \int_0^{\infty} d\omega' \omega^2 \left\{ \frac{(\bar{q}^2_\omega)^2}{2} + (\bar{q}^2_\omega q^2_\omega)^2 \right\} \right\} \right]$$

(96)

denoting $\omega_k$ simply as $\omega$.

By similar calculations one has

$$\bar{E} = \left( \frac{al}{4\pi} \right) \int_0^{\infty'} d\omega \left\{ \bar{q}^2_\omega + \omega^2 \bar{q}^2_\omega \right\} .$$

(97)

Because $\bar{q}^2_\omega = \omega^2 \bar{q}^2_\omega$ and the average values in the second integral of Eq. (96) are zero, it holds that for the quantities contained in the frequency interval $d\nu = d\omega/2\pi$

$$\bar{\Delta}^2 = \frac{(\bar{E})^2}{az} .$$

(98)

For the present case of the one-dimensional string, $a$ represents the volume of the system and $z_\nu$ is equal to the number of the eigen-oscillations contained in $d\nu$, namely two. This equation corresponds, therefore, to the fluctuation of the energy of radiation resulting from the Rayleigh-Jeans formula of radiation, which was demonstrated by Einstein.

Then, "by passing to quantum mechanics", $u$, $H$, $q_k$ and $E$ become matrices. But $x$ remains a number, because, if a series of mass points is considered instead of the string, $x$ "means the number of a particular definite mass point". The element of $q_k$, $q_k(nm)$, is zero except for

$$\left\{ \begin{array}{l} n_j - m_j = 0 \ (j \neq k) \\ n_k - m_k = \pm 1. \end{array} \right.$$

(99)

If the elements of a matrix are averaged over their phases, the matrix becomes diagonal matrix whose elements are identical with the diagonal elements of the original matrix. The diagonal elements of the matrix corresponding to

---

57)Here the integration formula \( \lim_{a \to \infty} \frac{1}{a} \int_{-a}^{a} \sin \omega_\alpha f(\omega)d\omega = \pi f(0) \) is used.

58) Cf. Sec. 3, §1, Chap. 1, Vol. II.

59) Born et al. use for $q_k$ the result obtained by them for a one-dimensional oscillator.
the first integral in Eq. (96) contribute, therefore, to \( \Delta^2 \) contained in the frequency interval \( dv \) by an amount \( (E^*)^2/2a \), where \( E^* \) is the sum of "the average thermal energy" \( \bar{E} \) and "the zero point energy" \( h\nu \).

On the other hand, from the relation

\[
-q_j\dot{q}_j(nn) = \dot{q}_j q_j(nn) = \frac{h}{2l \pi i}
\]

valid for a one-dimensional oscillator, one sees that the diagonal elements of the matrix corresponding to the second integral in Eq. (69) contribute to \( \Delta^2 \) by an amount \(- (1/2)(h\nu)^2/a\). Quantum mechanics gives, thus, instead of Eq. (98),

\[
\Delta^2 = h\nu \bar{E} + \frac{\bar{E}^2}{z \nu V}; \quad (V = a).
\]

This corresponds to the fluctuation of the energy of radiation resulting from the Planck formula of radiation.

As we have seen in Sec. 2, §4, Chap. 5, Vol. II, the Bose-Einstein statistics based on the theory of light quantum gives the Planck formula of radiation, and hence the fluctuation of the energy of radiation similar in character to Eq. (100). As we have seen also there, Einstein showed that this statistics gives the same kind of fluctuation of the density of a gas. Einstein adopted the de Broglie theory of wave in the molecular theory of gas, since the fluctuation of the density of a gas corresponding to the second term in Eq. (100) should come from the wave nature.

Born-Heisenberg-Jordan succeeded next in showing the existence of the first term in Eq. (100), which comes from the particle nature, by applying matrix mechanics to the wave motion, using the amplitudes of the eigen oscillations of the string as the canonical coordinates. As we have described before in Sec. 4, §1, Chap. 3 and Sec. 3, §1, Chap. 5 of Vol. II, Rubinowicz showed in 1917 that the energy of cavity radiation was given in the form \( \sum n_k h\nu_k \), by applying Sommerfeld’s quantum condition to each of the eigen oscillations which are superposed to give rise to cavity radiation. Rubinowicz’s work was, as we pointed out there, the germ of the quantization of wave field. Born-Heisenberg-Jordan’s matrix mechanics of the oscillation of a string was the beginning of the quantum theory of fields.

\(^{60}\) Note that \( z \nu \) = 2 in the present case.
§7. **Pauli’s derivation of the Balmer formula by matrix mechanics**

As our previous descriptions in the present chapter show, Heisenberg presented the "quantum-theoretical quantum condition" to replace the quantum condition in the old quantum mechanics, and Born and Jordan formulated it by introducing matrix mechanics as the "intensified quantum condition", which was then generalized into the "commutation relations" for a system of multiple degrees of freedom by them in collaboration with Heisenberg. The quantum condition in the old quantum mechanics was presented originally to grasp, in the framework of the classical mechanics, the existence of stationary states in the Bohr model of atom, and the explanation according to it of the Balmer formula. It was an important subject also in matrix mechanics to derive the Balmer formula.

This subject was accomplished by Pauli, in his paper received on Jan. 17, 1926 "On the Hydrogen Spectrum from the Standpoint of the new Quantum Mechanics". In a paper received five days later, namely on Jan. 22, 1926, Dirac treated hydrogen atom using his method of quantum algebra, as we shall describe in the next section. And, in a paper received further five days later, that is on Jan. 27, 1926, Schrödinger derived the Balmer formula as solution to the eigenvalue problem of the wave equation proposed by him for the motion of the de Broglie wave. The developments in the finale to arrive at the establishment of quantum mechanics became *allegro e vivace* in 1926.

Now, in his paper under consideration, Pauli makes use of the method that was applied by Lenz in a treatment, on the basis of the old quantum mechanics, of the splitting of the spectral lines of a hydrogen-like atom in mutually perpendicular magnetic and electric fields. Let the electric charge of the nucleus of a hydrogen-like atom be $Ze$, and the mass, position vector, momentum and angular momentum of the electron be, respectively, $m_0$, $\mathbf{r}$, $\mathbf{p}$ and $\mathbf{M}$. If $\mathbf{A}$ is defined to be the vector that directs from the nucleus to the aphelion of the elliptic orbit of the electron and has a length equal to the eccentricity of the orbit, it is given by

$$\mathbf{A} = \left( \frac{1}{Ze^2m_0} \right) \mathbf{M} \times \mathbf{p} + \frac{\mathbf{r}}{r}. \quad (101)$$

---

The scalar product $\vec{A}$ by $\vec{r}$,

$$\vec{A} \cdot \vec{r} = -\left(\frac{1}{Ze^2m_0}\right)\vec{M}^2 + \vec{r}, \quad (102)$$

represents the equation of a conic section. In terms of the energy of the system $E$, the square of $\vec{A}$ is given by

$$1 - \vec{A}^2 = -\left(\frac{2}{Z^2e^4m_0}\right)E\vec{M}^2. \quad (103)$$

These quantities in the classical mechanics are then matrixized, and are denoted in Gothic as $E, \vec{r}, \vec{p}, \vec{M}$ and $\vec{A}$. Because in the matrix mechanics $\vec{M} \times \vec{p} \neq -\vec{p} \times \vec{M}$ while in the classical mechanics $\vec{M} \times \vec{p} = -\vec{p} \times \vec{M}$, Pauli symmetrized them to define $\vec{A}$ as

$$\vec{A} = \left(\frac{1}{Ze^2m_0}\right)\left(\frac{\vec{M} \times \vec{p} - \vec{p} \times \vec{M}}{2}\right) + \frac{\vec{r}}{r}. \quad (104)$$

Making use of the commutation relations for the components $x, p_x$, etc. of $\vec{r}$ and $\vec{p}$, he derives the relations

$$\vec{M} \times \vec{M} = i\hbar\vec{M}, \quad (102-I)$$

and

$$\begin{cases} 
A_x M_z = M_x A_x, 
\ldots \\
A_y M_y - M_y A_x = M_x A_y - A_y M_x = i(\hbar/2\pi)\hbar A_x, \ldots \\
\vec{A} \cdot \vec{M} = \vec{M} \cdot \vec{A} = 0, 
\end{cases} \quad (102-II)$$

$$\vec{A} \times \vec{A} = -i\frac{\hbar}{2\pi}(\frac{2}{Z^2e^4m_0})E\vec{M}, \quad (102-III)$$

$$1 - \vec{A}^2 = -\left(\frac{2}{Z^2e^4m_0}\right)E\{\vec{M}^2 + (\hbar)^2\}. \quad (102-IV)$$

Equation (102-I) is the commutation relation for the angular momentum studied by Born-Heisenberg-Jordan. Pauli expresses the $(km, km)$ elements ($-k \leq m \leq k$) of the diagonalized $M_z$ and $\vec{M}^2$ as

$$M_{z,k,m} = m\hbar, \quad (\vec{M}^2)_{k,m} = \hbar^2(k+1). \quad (103a)$$
Letting $m$ change by $\pm 1$ to correspond to left and right circulations, respectively, in the $(x, y)$ plane, he uses the relation $M_{x_{k, m \pm 1}}^{k, m} = \pm i M_{y_{k, m \pm 1}}^{k, m}$ to get
\[
|M_{x_{k, m \pm 1}}^{k, m}|^2 = |M_{y_{k, m \pm 1}}^{k, m}|^2 = (\hbar/2\pi)^2(k \pm m)(k \mp m + 1). \quad (103b)
\]

Further, he shows that the elements of $A$ defined by
\[
\begin{align*}
A_{x_{k', m \pm 1}}^{k, m} &= \pm i A_{x_{k, m \pm 1}}^{k, m}; (k' = k \pm 1) \\
|A_{x_{k+1, m \pm 1}}^{k, m}|^2 &= |A_{y_{k+1, m \pm 1}}^{k, m}|^2 = C_k^{k+1}(k \mp m)(k \mp m + 1)/4,
\end{align*}
\]
where $C_k^{k+1}$ is a function of $k$ with the property $C_k^{k+1} = C_{k+1}^k (\geq 0)$, satisfy Eq. (102-II). With the use of these equations, Pauli then gets
\[
|A_{x_{k+1, m \pm 1}}^{k, m}|^2 = |A_{y_{k+1, m \pm 1}}^{k, m}|^2 = C_k^{k+1}(k \pm m + 1)(k \pm m + 2)/4, \quad (104b)
\]
\[
|A_{x_{k, m \pm 1}}^{k+1, m}|^2 = C_k^{k+1}\{(k + 1)^2 - m^2\}. \quad (104c)
\]

Because the application of Eqs. (102-I) and (102-II) to the commutation relation between $M_y$ (or $M_x$) and the $z$-component of Eq. (102-III) gives, on account of the diagonality of $E$, the $x$- (or $y$-) component of Eq. (102-III), it suffices to treat the $z$-component of Eq. (102-III). Pauli takes its $(km, km)$ element and makes use of Eqs. (103a) and (104), getting the following recurrence formula for $C_k^{k-1}$,
\[
m\{-2k + 3\}C_k^{k+1} + (2k - 1)C_k^{k} = \frac{|E|m}{R\hbar Z^2}, \quad (105)
\]
where $E$ is an diagonal element of $E$ $(E < 0)$ and $R$ is the Rydberg constant defined by $R = 2\pi^2 e^4 m_0/\hbar^3$.

Since for the possible minimum value of $k$ under a given $|E|$ one should put $C_k^{k-1} = 0$, the coefficient of $m$ in the left-hand side of Eq. (105) cannot be positive. On the other hand, the coefficient of $m$ in the right-hand side of Eq. (105) is positive. Therefore, for the minimum value of $k$, only the value $m = 0$ is permitted, so that the minimum value of $k$ is also zero. Thus, if the maximum value of $k$ permitted under a given $|E|$ is denoted by $n^*$, it follows that
\[
k = 0, 1, 2, \ldots, n^*. \quad (106)
\]
With the value $C_{n^*+1} = 0$ taken as a matter of course, one can use Eq. (105) in a recurrent way beginning with $k = n^*$ to determine $C_{n^*-1}, C_{n^*-2}, \ldots, C_0$.

From the values of $C_{k-1}^k$ so obtained and Eq. (104), one gets

$$(\vec{A})^{k,m}_{k,m} = \left(\frac{|E|}{RhZ^2}\right) \{n^*(n^* + 2) - k(k + 1)\}.$$  

Inserting this and Eq. (103a) into Eq. (102-IV), one then has

$$1 = \left(\frac{|E|}{RhZ^2}\right) (n^* + 1)^2.$$  

This gives, with the definition $n = n^* + 1$,

$$|E| = \frac{RhZ^2}{n^2},$$  

which is nothing but the Balmer formula.

In state with quantum number $n$, $k$ takes values $k = 0, 1, 2, \ldots, n - 1$. In the "normal state" with $n = 1$, "the quantum number $m$ can take no value other than $m = 0$, and therefore this state is nonmagnetic". Pauli says that "this conclusion may seem strange, particularly in analogy with the behavior of alkali atoms". In the old quantum mechanics, the normal (or ground) state of an alkali atom was considered to be the $1s$-state, that is the state with $n = k = 1$. States with $k = 0$ were excluded, because any orbit with $k = 0$ meant a linear motion going forth and back through the nucleus. The assignment of $k = 1$ brought about a magnetic moment to the ground state of an alkali atom, but could not explain the anomalous Zeeman effect. In Heisenberg's model in 1922, of which mention has been made in §4 of the present section, this $k = 1$ was divided between the valence electron and the nucleus.

Pauli considers, however, that "the modifications still necessary for the interpretation of the anomalous Zeeman effect, of the foundations of the new quantum mechanics could possibly become apparent even in the case of atoms with a single electron". He states that "the result that the normal state of such an atom should be nonmagnetic is probably not yet to be considered as conclusive", and refers to Goudsmidt-Uhlenbeck's "specific proposal". The application of matrix mechanics to Goudsmit-Uhlenbeck's model is done in Heisenberg-Jordan's paper treated in §4 of the present section, a paper received two months later than Pauli's paper under consideration.
Pauli makes a calculation of the change of energy of a hydrogen-atom, which is caused by an electric as well as a magnetic field imposed mutually perpendicular to each other. Treating the coexistent external electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$ as perturbations, he shows that the time average of the change of energy of the atom is given by the diagonal matrix

$$E_1 = \left( \frac{3ea}{2} \right) (\mathbf{E} \cdot \mathbf{A}) + \left( \frac{e}{2m_0c} \right) (\mathbf{H} \cdot \mathbf{M})$$

where $a = Ze^2/2|E|$. With the introduction of the notations

$$\left( \frac{2\pi}{\hbar} \right) \mathbf{M} = \mathbf{I}_1 + \mathbf{I}_2, \quad \sqrt{\frac{Z^2 R \hbar}{|E|}} \mathbf{A} = \mathbf{I}_1 - \mathbf{I}_2,$$

$$\tilde{\nu}_m = \frac{e\mathbf{H}}{4\pi m_0 c}, \quad \tilde{\nu}_e = \frac{(3/4\pi)e\mathbf{E}}{(2m_0|E|)^{1/2}},$$

$$\nu_1 = \nu_m + \nu_e, \quad \nu_2 = \nu_m - \nu_e,$$

this can be rewritten as

$$E_1 = (\mathbf{I}_1 \cdot \nu_1)\hbar + (\mathbf{I}_2 \cdot \nu_2)\hbar.$$

The magnitudes of $\nu_m$ and $\nu_e$ are equal, respectively, to the Larmor frequency and the frequency characteristic of the Stark effect. $\mathbf{I}_1$ and $\mathbf{I}_2$ are commutable each other, and satisfy the relations $\mathbf{I}_1 \times \mathbf{I}_1 = i\mathbf{I}_1$ and $\mathbf{I}_2 \times \mathbf{I}_2 = i\mathbf{I}_2$, just as for $\mathbf{M}$. Because $\mathbf{I}_1^2 = \mathbf{I}_2^2 = (n^*/2)(n^*/2 + 1), n^*/2$ plays here the role of the quantum number $k$ of $\mathbf{M}$. $E_1$ is diagonal when the component of $\mathbf{I}_1$ in the direction of $\nu_1$ and that of $\mathbf{I}_2$ in the direction of $\nu_2$ are diagonal. These situations are similar to the diagonalization of $\mathbf{M}^2$ and $\mathbf{M}_z$, such that if the quantum numbers of $\mathbf{I}_2$ and $\mathbf{I}_2$, corresponding to the quantum number $m$ of $\mathbf{M}$, are denoted by $n_1^*/2 - n_1$ and $n_2^*/2 - n_2$ respectively, then

$$E_1 = \left( \frac{n^*}{2} - n_1 \right) \nu_1 \hbar + \left( \frac{n^*}{2} - n_2 \right) \nu_2 \hbar,$$  \hspace{1cm} (108)$$

where $0 \leq n_1 \leq n_1^*$ and $0 \leq n_2 \leq n_2^*$.

In the old quantum mechanics, the quantum condition was imposed on the orbit of the electron calculated with the use of classical mechanics. In the case of a simultaneous application of magnetic and electric fields, the orbit could be transformed into a linear-segmental orbit passing through the nucleus by
changing adiabatically the relative directions as well as the strengths of both the fields. Because of this, the problem of the effect of crossing magnetic and electric fields brought about difficulties in the old quantum mechanics, in which orbits with \( k = 0 \) were excluded. Pauli states that “in the new quantum mechanics, in which any illustration of the stationary states by particular electron orbits is not introduced, special prohibitions become superfluous and the mentioned difficulties fall off of themselves”. Though this makes possible one of the results achieved in matrix mechanics, it remains unknown how to understand physically those motions with \( k = 0 \) that are derived from the angular momentum represented by a matrix.

3. Dirac’s Formulation by Quantum Algebra

§1. Poisson bracket expressions and commutation relations

Dirac formulated Heisenberg’s approach, independently of Born-Jordan, by the method of quantum algebra due to himself, in his paper “The Fundamental Equations of Quantum Mechanics”,\(^ {63}\) which was received on Nov. 7, 1925, about forty days later than Born-Jordan’s paper. It was nine days later than Dirac’s paper under consideration that Born-Heisenberg-Jordan’s paper mentioned in the preceding section was received. Dirac’s second paper on this subject was received further about one month later. In this way, the formulation by the method of matrices and that by the method of quantum algebra proceeded almost at the same time. In Dirac’s first paper mentioned above no reference was made to Born-Jordan’s work.

In the first paper, Dirac treats a system of \( u \) degrees of freedom from the beginning, differently from Born-Jordan who first studied a system of one degree of freedom. He expresses the development of each coordinate \( x \) of the system in the form of a multiple Fourier series as

\[
x = \sum_{\alpha_1 \cdots \alpha_u} x(\alpha_1 \alpha_2 \cdots \alpha_u) \exp\{i(\alpha_1 \omega_1 + \alpha_2 \omega_2 + \cdots + \alpha_u \omega_u)t\} \\
= \sum_{\alpha} x_\alpha \exp\{i(\alpha \omega)t\}. \tag{109}
\]

Since any solution to the set of equations of motion is specified with a set $\kappa$ of $u$ constants $\kappa_1, \ldots, \kappa_u, x_\alpha$ and $(\alpha \omega)$ are functions of $\alpha$ and $\kappa$, and may be written $x_{\alpha \kappa}$ and $(\alpha \omega)_{\kappa}$.

In quantum mechanics, "according to Heisenberg, we still assume that each coordinate can be represented by harmonic components of the form $\exp(i\omega t)$, the amplitude and frequency of each depending on two sets of numbers $n_1, \ldots, n_u$ and $m_1, \ldots, m_u$, in this case all integers, and being written as $x(nm)$, $\omega(nm)$". Dirac says that "the quantum solutions are all interlocked, and must be considered as a single whole". Following Heisenberg, the amplitude of the product $xy$ is determined by the equation

$$xy(nm) = \sum_k x(nk)y(km),$$

which Dirac calls "the Heisenberg product".

In Heisenberg’s theory, only differentiation with respect to the time $t$ was considered. Dirac then proceeds to "determine the form of the most general quantum operation $d/dv$” that satisfy the laws

$$\frac{d}{dv}(x + y) = \frac{d}{dv}x + \frac{d}{dv}y, \quad (111a)$$

$$\frac{d}{dv}(xy) = \frac{d}{dv}x \cdot y + x \frac{d}{dv}y. \quad (111b)$$

Equation (111a) requires the amplitudes of the components of $dx/dv$ to be linear functions of $x$, that is

$$\frac{dx}{dv}(nm) = \sum_{n'm'} a(nm; n'm')x(n'm'). \quad (112)$$

Equation (111b) imposes conditions on the coefficients $a$. With the use of Eq. (112) for Eq. (116) it follows that

$$\sum_{n'm'k} a(nm; n'm')x(n'k)y(km') = \sum_{kn'k'} a(nk; n'k')x(n'k')y(km) + \sum_{kk'm'} x(nk)a(km; k'm')y(k'm').$$

A comparison of the coefficients of $x(n'k)y(k'm)$ gives

$$\delta_{kk'}a(nn; n'm') = \delta_{nn'}a(nk'; n'k) + \delta_{nn'}a(km; k'm').$$
Applying this equation to the case of \( k = k', m \neq m', n \neq n' \), that of \( k \neq k', m = m', n \neq n' \), and so on, Dirac shows that

\[
\begin{align*}
a(nk; n'k') &= 0 \quad (n \neq n', k \neq k'), \\
a(nk'; nk) &= -a(km; k'm) = a(kk')(k \neq k'), \\
a(nk; nk) &= a(kk) - a(nn),
\end{align*}
\]

where \( a(kk') \) is a quantity dependent on the pair of integers \( k \) and \( k' \).

Inserting these expressions into Eq. (112), he gets

\[
\frac{dx}{dv}(nm) = \sum_k \{x(nk)a(km) - a(nk)x(km)\}. \quad (113)
\]

As the quantities in the right-hand side of this equation are Heisenberg products, this gives

\[
\frac{dx}{dv} = xa - ax. \quad (114)
\]

The frequency \( \omega(nm) \) is given in the form \( \omega(nm) = \Omega(n) - \Omega(m) \) according to Ritz's combination rule. Dirac calls \( \Omega(n) \) "the frequency level". If "\( a \) is a constant, so that \( a(nm) = 0 \) except when \( n = m \)", and "in particular, if \( ia(mm) = \Omega(m) \)", one has

\[
\frac{dx}{dv}(nm) = i\omega(nm)x(nm). 
\]

This is the "ordinary differentiation with respect to \( t \)."

Dirac then asks the question as to what the expression \( (xy - yx) \) corresponds on the classical theory. He supposes that \( x(n, n - \alpha) \) varies only slowly with the \( n \), the \( n \)'s being large numbers and the \( \alpha \)'s small ones, so that it is possible to put\(^64\)

\[
x(n, n - \alpha) = x_{\alpha \kappa}; \quad \kappa_r = n_r h. \quad (115)
\]

In this case, one has

\[
x(n, n - \alpha)y(n - \alpha, n - \alpha - \beta) - y(n, n - \beta)x(n - \beta, n - \alpha - \beta) \\
= \{x(n, n - \alpha) - x(n - \beta, n - \beta - \alpha)\}y(n - \alpha, n - \alpha - \beta)
\]

\(^64\)Dirac's treatment here is of the same kind as Heisenberg's for a system of one degree of freedom based on the correspondence principle.
Dirac's Formulation by Quantum Algebra

\[- \{y(n, n-\beta) - y(n-\alpha, n-\alpha-\beta)\} x(n-\beta, n-\alpha-\beta)\]

\[= \hbar \sum_r \left\{ \beta_r \left( \frac{\partial x_{\alpha \kappa}}{\partial \kappa_r} \right) y_{\beta \kappa} - \alpha_r \left( \frac{\partial y_{\beta \kappa}}{\partial \kappa_r} \right) x_{\alpha \kappa} \right\} . \]

Because with the angle variables \(w_r\) defined as \(w_r = \omega_r t/2\pi\), one has the relations

\[2\pi i\beta_r y_{\beta \kappa} \exp\{i(\beta \omega)t\} = \frac{\partial}{\partial w_r} [y_{\beta \kappa} \exp\{i(\beta \omega)t\}] , \]

it turns out that \((xy - yx)\) corresponds to

\[- \frac{i\hbar}{2\pi} \sum_r \left\{ \frac{\partial x}{\partial w_r} \frac{\partial y}{\partial w_r} - \frac{\partial y}{\partial w_r} \frac{\partial x}{\partial w_r} \right\} . \]

If \(\kappa_r\) is taken to be equal to the action variable \(J_r\), this expression becomes \(i\hbar/2\pi\) times the Poisson (or Jacobi) bracket

\[[x, y] = \sum_r \left\{ \frac{\partial x}{\partial w_r} \frac{\partial y}{\partial J_r} - \frac{\partial y}{\partial w_r} \frac{\partial x}{\partial J_r} \right\} . \]

It is known that \([x, y]\) is invariant, if the set of pairs of variables \(w_r\) and \(J_r\) is transformed to any set of the pairs of canonical variable \(q_r\) and \(p_r\).

The elementary Poisson bracket expressions are those for the \(q_r\)'s and \(p_r\)'s themselves, namely

\[\left\{ \begin{array}{c} [q_r, q_s] = [p_r, p_s] = 0 , \\ [q_r, p_s] = \delta_{rs} . \end{array} \right\} \tag{116} \]

If \(x\) and \(y\) are given as algebraic functions of \(p_r\) and \(q_r\), \([x, y]\) can be calculated from Eq. (116) with the use of the following relations,

\[\left\{ \begin{array}{c} [x + y, z] = [x, z] + [y, z] , \\ [xy, z] = [x, z]y + x[y, z] , [x, y] = -[y, x] . \end{array} \right\} \tag{117} \]

Dirac says that "the bracket expression \([x, y]\) thus has a meaning on the quantum theory when \(x\) and \(y\) are quantum variables, if we take the elementary bracket expressions to be still given by the above equations", and makes the following "fundamental assumption". That is, "the difference between the
Heisenberg products of two quantum quantities is equal to $i\hbar/2\pi$ times their Poisson bracket. In symbols,

$$xy - yx = \left( \frac{i\hbar}{2\pi} \right) [x, y].$$

(118)

He thinks that "it seems reasonable to take this equation as constituting the general quantum conditions". The only independent conditions given by Eq. (118) are

$$\begin{align*}
q_r q_s - q_s q_r &= 0, \\
p_r p_s - p_s p_r &= 0, \\
q_r p_s - p_s q_r &= \left( \frac{i\hbar}{2\pi} \right) \delta_{rs}. \\
\end{align*}$$

(119)

In the classical theory, the equations of motion are given in terms of the Hamiltonian function $H$ by

$$\dot{p}_r = [p_r, H], \quad \dot{q}_r = [q_r, H].$$

Dirac thinks that "they may be taken to be true" on quantum theory for any $x$:

$$\dot{x} = [x, H].$$

(120)

It results from Eqs. (118) and (120) given by Dirac that

$$\dot{x} = \left( \frac{2\pi i}{\hbar} \right) (Hx - xH).$$

This is equivalent to Eq. (64) in the preceding section, which was given by Born-Jordan. Equation (119) given by Dirac is equivalent to the commutation relations, Eq. (69) in the preceding section, which was given by Born-Jordan first for a system of one degree of freedom, and then generalized by them in collaboration with Heisenberg to a system of many degrees of freedom.

§2. $q$-numbers and $c$-numbers — quantum algebra

Dirac says, in his second paper received on Jan. 22, 1926, "Quantum Mechanics and a Preliminary Investigation of the Hydrogen Atom", that because the variables used for describing a dynamical system are not commutable as is seen from the Heisenberg product, "they are not numbers in the sense of the word previously used in mathematics", and calls them $q$-numbers. The numbers of

\[^{65}\text{P. A. M. Dirac, Proc. Roy. Soc. A110 (1926), 561.}\]
classical mechanics are, on the other hand, called c-numbers. He thinks that "at present one can form no picture of what a q-number is like". He tries thus to derive conversely the premises taken in his first paper by assuming the conclusions for q-numbers arrived at in the first paper as the premises.

Namely, Dirac defines the Poisson bracket expression \([x, y]\) for the q-numbers \(x\) and \(y\) by Eq. (118), and assumes that "the elementary P.B.'s (Poisson brackets) are given" by Eq. (116) because of Eq. (118). If \(x\) and \(y\) are given functions of the \(p\)'s and \(q\)'s, the P.B. \([x, y]\) can be expressed in terms of the elementary P.B.'s by successive applications of Eq. (117). For example, if the function \(f(q_r)\) does not involve any number that does not commute with \(q_r\), one has

\[
[f(q_r), q_r] = \frac{\partial f}{\partial q_r}.
\]

By making use of this equation and the notation \((\alpha q) = \sum_r \alpha_r q_r\), in which \(\alpha_r\), \((r = 1, \ldots, u)\) are c-numbers, one can show the following relations,

\[
\begin{align*}
& e^{i(\alpha q)} f(q_r, p_r) = f(q_r, p_r - \alpha_r \hbar) e^{i(\alpha q)}, \\
& f(q_r, p_r) e^{i(\alpha q)} = e^{i(\alpha q)} f(q_r, p_r + \alpha_r \hbar).
\end{align*}
\]

If one effects the "transformation" of the variables \(q_r, p_r\) into the variables \(Q_r, P_r\) by the relations

\[
Q_r = bq_r b^{-1}, \quad P_r = bp_r b^{-1},
\]

where \(b\) is a q-number, \(Q_r, P_r\) are "canonical variables" that satisfy the elementary P.B.'s. Every kind of formula valid for \(q_r, p_r\) is therefore valid for \(Q_r, P_r\) too. Dirac says, however, that "these formulae do not appear to be of great practical value". A summary of the quantum algebra is given in his paper "On Quantum Algebra".

About the dynamical system itself, Dirac assumes that the equations of motion are given by Eq. (120) in terms of the g-number Hamiltonian \(H\). Now, "in order to able to get results comparable with experiment from our theory, we must have some way of representing q-numbers by means of c-numbers".

---

66) A function \(f(z)\) of a q-number \(z\) "can be defined only by an algebraic relation connecting \(f(z)\) with \(z\).

67) In the paper under consideration, Dirac denotes \(\hbar/2\pi\) simply as \(\hbar\). In our description, however, we use the present-day symbol \(\hbar = \hbar/2\pi\) to avoid confusion.

68) P. A. M. Dirac, Cambridge Phil. Soc. 23 (1926), 412 (received on Jul. 17, 1926).
Dirac studies this problem for the case of multiply periodic systems. He assumes that on the quantum theory a dynamical system is multiply periodic, when there exists a set of uniformising variables $J_r, w_r, (r = 1, \ldots, u)$ such that (i) $J_r, w_r$ are canonical variables, (ii) $H$ is a function of $J$’s only, and (iii) each of the original $p$’s and $q$’s can be expanded in either of the forms

$$\sum c_\alpha \exp\{i(\alpha w)\} = \sum \exp\{i(\alpha w)\}c'_\alpha, \quad (124)$$

where $(\alpha w) = \sum r_\alpha w_r$, the $\alpha_r$’s being integers, and $c_\alpha$’s and $c'_\alpha$’s are functions of $J_r$’s only.

For a multiply periodic system, it is supposed further that a function $x$ of the $q$’s and $p$’s can be expanded in a similar form

$$x = \sum x_\alpha \exp\{i(\alpha w)\} = \sum \exp\{i(\alpha w)\}x'_\alpha, \quad (125)$$

the $x_\alpha$’s and $x'_\alpha$’s being also functions of $J_r$’s only. Applying Eq. (122) with $q_r, p_r$ replaced by $w_r, J_r$ to this equation, one gets $x'_\alpha(J_r) = x_\alpha(J_r + \alpha_r \hbar)$. A comparison of Eq. (125) with its complex conjugate, gives $x_\alpha(J_r + \alpha_r \hbar) = x_{-\alpha}(J_r)$. "This relation is brought out more clearly", if the notation $x_\alpha(J_r)$ is changed to $x(J, J - \alpha \hbar)$. That is, instead of Eq. (125), one has

$$x = \sum x(J, J - \alpha \hbar)e^{i(\alpha w)} = \sum e^{i(\alpha w)}x(J + \alpha_r \hbar, J). \quad (126)$$

Taking the same kind of expansion for $y$ and calculating the product $xy$, one gets

$$xy(J, J - \gamma \hbar) = \sum x(J, J - \alpha \hbar)y(J - \alpha \hbar, J - \gamma \hbar) \quad (127)$$

with the use of Eq. (122). Dirac denotes the c-number obtained by substituting "c-number $n, \hbar$" for each $J_r$ in $x(J, J - \alpha \hbar)$ with $x(n, n - \alpha)$. The application of the equations of motion and Eq. (122) to $e^{i(\alpha w)}$ gives

$$\frac{de^{i(\alpha w)}}{dt} = (i/\hbar)\{H(J_r) - H(J_r - \alpha_r \hbar)\}e^{i(\alpha w)}. \quad (128)$$

69) The "complex conjugate" of a q-number is defined as that obtained by inverting the order of the q-number factors in the q-number under consideration and changing $i$ to $\bar{i}$. A q-number is called a "real q-number" when it is equal to its complex conjugate. The elemental P.B.'s given in Eq. (119) are invariant for transformation into their complex conjugates. $x$ and $J_r$ are assumed to be real q-numbers.
The coefficient of $e^{i(\alpha \omega)}$ in the right-hand side of this equation is rewritten, similarly to the case of $x_\alpha (J_r)$, as $i\omega (J, J - \alpha \hbar)$. The use of the $c$-numbers mentioned above brings the equations of motion into the form

$$\dot{x}(n, n - \alpha) = i\omega (n, n - \alpha) x(n, n - \alpha),$$

and Eq. (127) into the form

$$xy(n, n - \gamma) = \sum x(n, n - \alpha) y(n - \alpha, n - \gamma).$$

These two equations are those presumed in the first paper.

Dirac concludes, thus, “we may consider the aggregate of all the $c$-numbers, $x(n, n - \alpha) = \exp \{i\omega (n, n - \alpha)t\}$, as representing the values of the $q$-number $x$ for all values of the $q$-numbers $J_r$”. Namely, it is shown that the $q$-number corresponding to any physical quantity in a multiply periodic system has “representations” which are equal to the matrix elements of the quantity in matrix mechanics. Dirac says, however, that “a $q$-number still has a meaning and can be used in the analysis when it is not multiply periodic, although there is at present no way of representing it by $c$-numbers”.

§3. Application of the quantum algebra

§3.1 Electron orbit in hydrogen atom

In the second paper under consideration, Dirac applies the algebra of $q$-numbers to “orbital motion in hydrogen atom”. Taking the $(x, y)$ coordinates in the orbital plane, he assumes the Hamiltonian to be

$$H = \left(\frac{1}{2m}\right)(p_x^2 + p_y^2) - \frac{e^2}{(x^2 + y^2)^{1/2}},$$

as in classical theory. Transforming the Cartesian coordinates $(x, y)$ into polar coordinates $(r, \theta)$, he rewrites it as

$$p_r = \frac{1}{2}(p_x \cos \theta + \cos \theta p_x) + \frac{1}{2}(p_y \sin \theta + \sin \theta p_y),$$

$$k = xp_y - yp_x.$$

---

70) About this point the reader is referred to Sec. 3, §2, i), Chap. 3 of the present volume.

71) $\cos \theta$ and $\sin \theta$ are defined by $\exp (i\theta) = \cos \theta + i\sin \theta$. In the first equation of Eq. (132), $p_x$ and $\cos \theta$, as well as $p_y$ and $\sin \theta$, are symmetrized because they are $q$-numbers.
It can be shown that \((r, \theta)\) and \((p_r, k)\) are pairs of canonical variables, by means of Eq. (121). \(p_x\) and \(p_y\) are expressed in terms of \(p_r\) and \(k\) as

\[
\begin{aligned}
 p_x + ip_y &= \left(p_r + \frac{ik_2}{r}\right) e^{i\theta} = e^{i\theta} \left(p_r + \frac{ik_1}{r}\right), \\
p_x - ip_y &= \left(p_r - \frac{ik_1}{r}\right) e^{-i\theta} = e^{-i\theta} \left(p_r - \frac{ik_2}{r}\right),
\end{aligned}
\]

(133)

where \(k_1 = k + \hbar/2\) and \(k_2 = k - \hbar/2\), so that one has \(k_2 e^{i\theta} = e^{i\theta} k_1\) and \(k_1 e^{-i\theta} = e^{-i\theta} k_2\) from Eq. (122).

The last two relations together with the relation \(p_x^2 + p_y^2 = (p_x - ip_y)(p_x + ip_y)\) then give

\[
H = \left(\frac{1}{2m}\right) \left(p_r^2 + \frac{k_1 k_2}{r^2}\right) - \frac{e^2}{r}.
\]

(134)

However, if we had originally assumed that the Hamiltonian was the same function of the polar variables as in classical theory, we should have had, instead,

\[
H = \left(\frac{1}{2m}\right) \left(p_r^2 + \frac{k^2}{r^2}\right) - \frac{e^2}{r}.
\]

(135)

Dirac says that “the only way to decide which of these assumptions is correct is to work out the consequences of both and see which agrees with experiment”.

Corresponding to the classical equation of elliptic motion \(l/r = 1 + \epsilon \cos(\theta - \alpha)\), Dirac puts

\[
l/r = a_0 + a_1 e^{i\theta} + a_2 e^{-i\theta}.\]

(136)

That is, he tries to find the equation of orbit in \(q\)-number. When Eq. (134) is used for \(H\), the \(p_r\) which is obtained from the equation of motion for \(e^{i\theta}\) and the time derivative of Eq. (136) coincides with the \(p_r\) which is obtained from the equation of motion for \(p_r\) itself, if \(a_0\) is put equal to \(m e^2/k_1 k_2\). On the other hand, in order to get the same kind of coincidence as the above with the use of Eq. (135) for \(H\), \(k' = (k^2 + \hbar^2/4)^{1/2}\) and \(\theta' = \theta k'/k\) should be used in place of \(k\) and \(\theta\). These represent an ellipse with a rotating apse line. The Hamiltonian given by Eq. (135) “is thus untenable”.

Applying the relations \(k_1/r + ip_r = me^2/k_1 + 2ka_1 e^{i\theta}\), \(k_1/r - ip_r = me^2/k_1 + 2e^{-i\theta} ka_2\) and others to “the constants of integration \(a_1\) and \(a_2\)”, which are \(q\)-numbers, to derive \((k_1/r + ip_r)(k_1/r - ip_r) = 2m(H + e^2/r)\), Dirac shows that
$H$ can be given in the form

$$H = -\frac{me^4}{2P^2}$$

(137)

in terms of a quantity $P$ which commutes with $k, a_1$ and $a_2$. The quantum-theoretical equation of the orbit turns out to have the form

$$\frac{1}{r} = \frac{me^2}{k_1k_2} \left( 1 + \frac{k_2}{2k} \varepsilon_1 e^{-i\theta} e^{i\theta} + \frac{k_1}{2k} \varepsilon_2 e^{i\theta} e^{-i\theta} \right),$$

(138)

where $\varepsilon_j$ is defined by $\varepsilon_j = \{1 - (k_j/P)^2\}^{1/2}, (j = 1, 2)$.

On the other hand, if one calculates $\dot{\theta}$ by considering that, in the Fourier expansion $\theta = w + \sum b_n e^{inw}, e^{inw}$ is expanded into a Fourier series in $\theta$, one finds that the constant term in the Fourier expansion of $r^2$ in $\theta$ is $k/mw$. Because it is not easy to calculate this constant term from Eq. (138), Dirac makes some lengthy indirect calculations to get the result

$$\dot{w} = \frac{me^4}{P^3}.$$ 

(139)

This is identical with $\partial H/\partial P$ obtained from Eq. (138), showing that $P$ is the action variable conjugate to $w$. Dirac says, thus, that “one would have to put $P$ equal to an integral multiple of $\hbar$”. In this way, Dirac fails to deduce the integer nature of $P$ but has to borrow the result of old quantum mechanics.

Dirac’s result, obtained by means of complicated calculations covering ten pages of the second paper, thus did not reach Pauli’s result described in §6 of the preceding section. This may be said to be due to the fact that Dirac attached so much importance to the correspondence between the classical and quantum Poisson brackets that he tried to derive a $q$-number equation for the orbit corresponding to elliptic orbit in the classical theory. In fact, Dirac emphasized with italics in the first paper that “the correspondence between quantum and classical theories lies not so much in the limiting agreement when $\hbar \to 0$ as in the fact that the mathematical operations on the two theories obey in many cases the same laws”.

Different from Dirac, Pauli did “not consider that the stationary states are represented by particular orbits”, as we have quoted before, but attached importance to the connection between the energy and angular momentum of the electron. However, Dirac’s idea that the matrix elements in matrix mechanics represent $q$-number dynamical variables plays a role in the process of establishment of quantum mechanics, as will be seen in Chap. 3 of the present volume.
§3.2 **Angular momentum and its canonical conjugate coordinate**

The application of the quantum algebra to angular momentum is carried out in Dirac’s third paper received on Mar. 27, 1926, “The Elimination of the Nodes in Quantum Mechanics”. Dirac mentions in a footnote to this paper, that “the angular momentum relations have been obtained independently by Born, Heisenberg, and Jordan”. The elimination of the nodes indicates a method in analytical mechanics which consists in obtaining a contact transformation from the Cartesian coordinates and momenta to a set of canonical variables, of which all except three are independent of the orientation of the system as a whole, while these three determine the orientation. Dirac says that the treatment of the dynamical problem of a number of particles moving in a central field of force and disturbing one another “begins by making this initial simplification”.

To do this in a quantum-mechanical way, Dirac “defines” $r$ and $p_r$ for an electron by the relations

\[ r = (x^2 + y^2 + z^2)^{1/2}, \quad rp_r = xp_x + yp_y + zp_z - i\hbar, \]  \hspace{1cm} (140)

where $p_r$ satisfies $[r, p_r] = 1$. He calculates P.B.’s between the square of the angular momentum $m^2$ and each of its components $m_x, m_y$ and $m_z$, and shows that

\[ p_x^2 + p_y^2 + p_z^2 = p_r^2 + \frac{m^2}{r^2}. \]  \hspace{1cm} (141)

One of the action variables to be introduced in classical mechanics is equal to $m$. But, since in quantum mechanics it “may not be equal to $m$”, Dirac denotes it by $k$. In the classical theory, a coordinate, say $z$, satisfies the relation $\frac{\partial^2}{\partial z^2} = -z$ together with the angle variable $\theta$, which is canonically conjugate to $m$. Expressed in P.B.’s, this becomes $[m, [m, z]] = -z$. Dirac tries “to choose our quantum variable so as to satisfy” the equation

\[ [k, [k, z]] = -z. \]  \hspace{1cm} (142)

From a comparison of $[k^2, [k^2, z]]$ with $[m^2, [m^2, z]]$, he gets

\[ m^2 = k^2 - (\hbar/2)^2. \]  \hspace{1cm} (143)

---

73) The results of his calculation correspond with the commutation relations for these quantities in matrix mechanics as mentioned in the preceding section.
Because Eq. (142) does not necessarily follow from the equation for \([k^2, [k^2, z]],\) he assumes further

\[ [k, [k, x]] = -x, \quad [k, [k, y]] = -y, \tag{144} \]

together with Eq. (142).\(^7\)

In the classical theory, any angle variable \(w\) can be expressed in the form

\[ e^{iw} = \left\{ \frac{a + ib}{a - ib} \right\}^{1/2} \]

in terms of real numbers \(a\) and \(b\). Since in the quantum theory the order of \(q\)-numbers matters, Dirac considers the "the correct quantum generalization" on this relation should be the more symmetrical

\[ e^{iw}(a - ib)e^{iw} = a + ib. \tag{145} \]

If \(J\) is an action variable such that \([J, a] = b\) and \([J, b] = -a\), one has the relation \([e^{iw}, J] = ie^{iw}\). From this relation it does no follow rigorously that \([w, J] = 1\), but Dirac says that this relation is sufficient for us to take \(w\) to be the variable conjugate to \(J\), since \(w\) occurs in the analysis only through \(e^{iw}\).

In consideration of Eq. (142) and the commutability of \(z\) with the \(z\)-component of the total angular momentum of the system \(M_z\), Dirac defines the angle variable \(\theta\) canonically conjugate to \(k\) by

\[ e^{i\theta}(z - i[k, z])e^{i\theta} = z + i[k, z], \tag{146} \]

which follows from Eq. (145) by putting \(a = z\) and \(b = [k, z]\). He defines also the angle variable \(\phi\) canonically conjugate to \(M_x\) by

\[ \phi e^{i\phi}(M_x - iM_y)e^{i\phi} = M_x + iM_y, \tag{147} \]

taking into account that \([M_z, M_x] = M_y, [M_z, M_y] = -M_x,\) and \(M_x\) commutes with \(k\) and \(j\) which is defined by \(M^2 = j^2 - (\hbar/2)^2\), a relation similar to Eq. (143).

By means of calculations with the use of quantum algebra for these quantities, Dirac shows that for a system of a single electron \(r, p, k, \theta; p (= m_z)\), are the new canonical variables, and the original coordinates \(x, y, z\) are connected

\(^7\)Dirac notes that Born-Heisenberg-Jordan state that negative values of \(k\) can be ignored without loss of generality, but this is justifiable only if it can be shown that transitions from a positive to a negative \(k\) cannot occur, and Eqs. (142) and (144) supply the necessary further assumption.
to them by the transformation equations

\[ x \pm iy = \frac{r}{2} \left\{ \frac{(k \pm p - \hbar/2)^{1/2}(k \pm p - 3\hbar/2)^{1/2}}{k^{1/2}(k - \hbar)^{1/2}} e^{i(\theta \pm \phi)} + \frac{(k \pm p + \hbar/2)^{1/2}(k \pm p + 3\hbar/2)^{1/2}}{k^{1/2}(k + \hbar)^{1/2}} e^{\pm i(\theta \pm \phi)} \right\} , \]  

(148)

\[ z = \frac{r}{2} \left\{ \frac{(k + p - \hbar/2)^{1/2}(k - p - \hbar/2)^{1/2}}{k^{1/2}(k - \hbar)^{1/2}} e^{i\theta} - \frac{(k + p + \hbar/2)^{1/2}(k - p + \hbar/2)^{1/2}}{k^{1/2}(k + \hbar)^{1/2}} e^{-i\theta} \right\} . \]  

(149)

For a system with two electrons, Dirac takes \( r, p_r, r', p_r'; k, \theta; k', \theta'; p, \phi; j, \psi \) as the set of the canonical variables, distinguishing the quantities related to the second electron by a prime. \( \phi \) is the same as the one defined by Eq. (147), but \( \theta \) is defined by

\[ e^{i\theta} (\vec{q} \cdot \vec{m'} - i[k, \vec{q} \cdot \vec{m'}]) e^{-i\theta} = \vec{q} \cdot \vec{m'} + i[k, \vec{q} \cdot \vec{m'}] , \]  

(150)

where \( \vec{q} = (x, y, z) \) and \( \vec{m'} = (m'_x, m'_y, m'_z) \). This definition is taken because of the fact that \([k, [k, \vec{q} \cdot \vec{m'}]] = -\vec{q} \cdot \vec{m'}\), and \( \vec{q} \cdot \vec{m} \) commutes with \( k', p, j, \theta' \) of the second electron is given by Eq. (150) with \( \theta, \vec{q}, k, \vec{m} \) replaced by \( \theta', \vec{q}', k', \vec{m'} \), respectively.

In order to define \( \psi \) canonically conjugate to \( j \), Dirac makes use of the relations satisfied by \( \vec{\mu} = \vec{m} \times \vec{m'} \) and \( \vec{M} \), namely, \( \vec{\mu} \cdot \vec{M} = 0 \), \([\mu_z, M_x] = \mu_y, \) \([\mu_z, M_y] = -\mu_x \) and \([\mu_z, M_z] = 0 \), which are of the same form as the relations between \( \vec{q} \) and \( \vec{m} \). In consideration of the commutability of \( \mu_z \) with \( r, p_r, r', p'_r, k, k' \) and \( p \), he defines \( \theta \) by the equation

\[ e^{i\theta} (\mu_z - i[j, \mu_z]) e^{-i\theta} = \mu_z + i[j, \mu_z] , \]  

(151)

which is of the same from as Eq. (146). For a system with more than two electrons, it suffices to apply the same procedure as given above to the angular momenta of the outermost electron and the atomic core.

Since, for the system of a single electron, one has \( m_x \pm im_y = \{k^2 - (p \mp \hbar/2)^2\}^{1/2} e^{\pm i\phi} \) in terms of \( \phi \) given above, it follows that \( k = \pm p \) are the limits of the values taken by the action variable \( k \), and therefore \( p \) takes \( 2|k| \) values between \( \pm p(|k| - \hbar/2) \) for a given value of \( k \). Dirac says that "it may be shown that \( k \) takes half integral quantum values when the central field consists of an
inverse square field of force with a small inverse cube field of force superposed”, and “there will thus be 1, 3, 5, \ldots stationary states for S, P, D, \ldots terms”.

For a system with two or more electrons, it can also be shown that \( j \pm p = 0 \) are boundary values for \( j \) and \( p \), and \( k \pm k' \pm j = 0 \) are boundary values for \( k, k' \) and \( j \). Hence \( p \) takes the values between \( \pm (|j| - h/2) \) and \( j \) takes the values, when \( k \) and \( k' \) are positive, between \( |k + k'| - h/2 \) and \( |k - k'| + h/2 \). “This rule applies generally for the addition of any two angular momenta”. With the use of these results, Dirac derives Landé’s \( g \)-factor by adopting “the usual model of the atom”, in which the atomic core is assumed to have the anomalous ratio of magnetic moment. No mention is made of Goudsmit–Uhlenbeck’s model.

In this way, different from Born-Heisenberg-Jordan who studied “the quantum mechanics” of angular momenta with the use of matrices, Dirac treated “the quantum mechanics” of angular variables themselves canonically conjugate to angular momenta, by attaching importance to the correspondence between classical and quantum Poisson brackets. With the presentation of the Schrödinger wave equation, it turned out later that the momentum canonically conjugate to the coordinate \( x \) was given by \( p_x = -i\hbar \partial / \partial x \), and a way became open to calculate much more easily the action variables canonically conjugate to angle variables, as differential operators, inversely in terms of angle variables.

4. Attempts at the Interpretation of Matrix Mechanics

§1. Lanczos’s attempt by means of integral equation

As we have described so far in the present chapter, following Heisenberg who grasped the atom as having the attributes of “aggregates of amplitudes”, Born and Jordan proposed matrix mechanics by representing each aggregate of amplitudes by a matrix, and developed it further in collaboration with Heisenberg, opening the way to the derivation by Pauli of the Balmer formula. However, the fundamental idea in matrix mechanics that physical quantities should be represented by matrices was strange to physics before that time. There arose, thus, attempts at interpreting matrix mechanics in terms of more familiar methods.

Lanczos attempted, in his paper received on Dec. 22, 1925 “On a Representation of the New Mechanics according to Field”,\(^{75}\) to transform “the

\(^{75}\)K. Lanczos, Zeit. f. Phys. 35 (1926), 812.
discontinuous view" in matrix mechanics into "a continuos view", in relation to the theory of integral equation. Lanczos's paper was received about one month later than Born-Heisenberg-Jordan's paper, but his discussions were based on the contents of Heisenberg's and Born-Jordan's papers.

Lanczos says in his paper that "Born and Jordan succeeded in giving Heisenberg's conception adequate mathematical expression in greatly intended generalization, and in revealing the general formal bases of the new theory". "One reaches a consistently built discontinuous theory, for which the classical concepts are merely in the sense of correspondence and as heuristic signposts of meaning, but which goes its particular way from the outset and connects a completely new meaning to the old symbol".

Lanczos then states that "an extraordinary simple and beautiful relation will be pointed out in the following, which is to be established between Heisenberg-Born-Jordan's theory and the theory of integral equations". "All results of the former theory can be expressed also in the form of integral equations, and therefore find out a formulation which may seem more natural than representation in matrix form to the usual physicists working with analytical methods". "At the same time there thus appears also a continuous view, which, as far as the description of facts is concerned, can stand with equal rights by the discontinuous view, since between the two views there exists a clear correspondence". Lanczos discussed, however, that "as far as the meaning of facts and hence the real essence of quanta are concerned, it is not excluded that the integral formulation is even superior to the matrix formulation, provided that the former has the merit to be set up really on itself, in immediate compatibility with the idea of field".

Now, let an eigen function of the integral equation with a nondegenerate symmetric kernel $K(s,\sigma) = K(\sigma, s)$ be denoted by $\varphi^i(s)$. The $\varphi^i(s)$'s are supposed to form a set of orthonormal functions, and $s$ or $\sigma$ denote a set of coordinates in a finite closed domain of finite dimension. Let a function of two points in this domain which is continuous at least stepwise be denoted by $f(s,\sigma)$. It is possible to expand $f(s,\sigma)$ as

$$f(s,\sigma) = \sum \sum a_{ik} \varphi^i(s) \varphi^k(\sigma),$$

with the expansion coefficient $a_{ik}$ given by

$$a_{ik} = \int f(s,\sigma) \varphi^i(s) \varphi^k(\sigma) dsd\sigma.$$
Attempts at the Interpretation of Matrix Mechanics

It results from Eq. (152) that

$$\int f(s, s)ds = \sum \sum a_{ik} \int \varphi^i(s)\varphi^k(s)ds = \sum a_{ii}. \quad (154)$$

For two functions $f(s, \sigma)$ and $g(s, \sigma)$, Lanczos defines $h(s, \sigma)$ as

$$h(s, \sigma) = \int f(s, \tau)g(\tau, \sigma)d\tau \equiv fg(s, \sigma), \quad (155)$$

and calls it “the field product” of $f$ and $g$. With the expansion coefficients $b_{mk}$ and $c_{ik}$ of $g$ and $h$, respectively, one has from the orthonormality of the eigen functions

$$c_{ik} = \sum_m a_{im}b_{mk}. \quad (154')$$

Equation (154) gives the diagonal sum of the matrix whose elements are the $a_{ik}$'s, and Eq. (154') the product of the matrices whose elements are the $a_{ik}$'s and $b_{mk}$'s.

Lanczos then considers the “field product” of many functions

$p, q, r, t, \ldots, u, v$

$$P(s, \sigma) = (pqrst \cdots uv)(s, \sigma). \quad (156)$$

The value of $\int P(s, s)ds$ is unchanged when the order of the factors in $P(s, \sigma)$ is altered cyclically, for example as $r, t, \ldots, u, v, p, q$. The variation of $\int P(s, s)ds$ with respect to the variation $\delta r$ of $r$, for example, is given by

$$\delta \int P(s, s)ds = \int \delta r(s, \tau)(t \cdots upq)(\tau, s)d\tau ds. \quad \delta \int P(s, s)ds = \int \delta r(s, \tau)(t \cdots upq)(\tau, s)d\tau ds. \quad (157)$$

This variation can be expressed as

$$\left\{ \begin{array}{l}
\delta \int P(s, s)ds = \int \delta r \frac{\partial P}{\partial r}(s, s)ds, \\
\frac{\partial P}{\partial r}(s, \sigma) = (t \cdots upq)(s, \sigma).
\end{array} \right. \quad (157)$$

Lanczos compares this “formation law” of differentiation with that given by Born-Jordan for the corresponding matrices.
Because the eigen function \( \varphi^k(\sigma) \) satisfies
\[
\int \varphi^k(\tau)K(\tau,\sigma)d\tau = \lambda_k \varphi^k(\sigma)
\]
with its eigen value \( \lambda_k \), one gets
\[
\int f(s,\tau)K(\tau,\sigma)d\tau = \sum (a_{ik}/\lambda_k)\varphi^i(s)\varphi^k(\sigma)
\]
by making use of Eq. (152). Interchanging the order of \( f \) and \( K \) in this equation one has
\[
\int K(s,\tau)f(\tau,\sigma)d\tau = \sum (a_{ik}/\lambda_k)\varphi^i(s)\varphi^k(\sigma).
\]
Hence if the “pointing (Punktierung)” of a function \( f \) is defined by
\[
\dot{f}(s,\sigma) = (Kf - fK)(s,\sigma),
\]
the expansion coefficients of \( \dot{f} \) are given by \((\lambda_i^{-1} - \lambda_k^{-1})a_{ik}\)’s. Now, in Born-Jordan’s theory the time derivative of \( a_{ik} \) is given by \( 2\pi i(W_i/h - W_k/h)a \), in terms of the level energies \( W_i \) and \( W_k \). If we “leave out” the factor \( 2\pi i \), and “define by our pointing the derivative, so to speak, with respect to \( 2\pi it \) instead of \( t \)”, it results that
\[
\frac{1}{\lambda_i} = \frac{W_i}{h}.
\]
Lanczos thus suggests that the inverse of each eigen value of the symmetric kernel \( K(s,\sigma) \) corresponds to a level energy.

He then makes the demand \( \delta I = 0 \) of the quantity
\[
I = \int (p\dot{q} - H(p, q))(s, s)ds
\]
for variations in \( q \) and \( p \), where \( H(p, q)(s, \sigma) \) denotes the “Hamiltonian” given for a system. From this demand one gets by making use of Eqs. (157) and (158),
\[
(Kq - qK)(s, \sigma) = \frac{\partial H}{\partial p}(s, \sigma), \quad (Kp - pK)(s, \sigma) = -\frac{\partial H}{\partial q}(s, \sigma).
\]
Lanczos calls these equations “determinative integral equations of \( p(s,\sigma) \) and \( q(s,\sigma) \)”. He assumes that the “Born-Jordan’s quantum condition” \( pq - qp = (h/2\pi i)1 \) corresponds to the integral equation
\[
(pq - qp)(s, \sigma) = HE(s, \sigma),
\]
(160)
where $E(s, \sigma)$ is the unit kernel given by

$$E(s, \sigma) = \sum \varphi^*(s) \varphi^*(\sigma),$$

(161)

which is zero for $s \neq \sigma$ and satisfies the condition

$$\lim_{\varepsilon \to 0} \int E(s, s + \varepsilon) d\varepsilon = 1.$$ 

One has therefore

$$(pq - qp)(s, \sigma) = \begin{cases} 0 & (s \neq \sigma) \\ \infty & (s = \sigma) \end{cases}.$$ 

(162)

Lanczos says the “if the functions $p(s, \sigma)$ and $q(s, \sigma)$ are to remain finite everywhere, Born-Jordan’s quantum condition can be valid only with arbitrary approximation, but not with full sharpness”.

Lanczos discussed in a footnote in his paper under consideration that, if Eq. (158) was supposed to correspond to Born-Jordan’s equation of motion, there would result the relation $H = \hbar K$. This is, however, in contradiction with his idea mentioned above, that each eigen value of $K$ would be inversely proportional to a level energy. This point was noted in a footnote in Schrödinger’s paper on the relation between matrix mechanics and wave mechanics, a paper which will be considered in the next chapter. Apart from this point, it should be noted that Lanczos treated the discontinuity which was supposed in matrix mechanics for the existence of quantum numbers indicating matrix elements, and attempted to connect it with the existence of discrete eigen values of an integral equation of continuous functions.

In Lanczos’s paper, however, no discussion was given of the meaning which the eigen functions of the integral equation would imply physically. As we shall see in detail in the next chapter, Schrödinger proposed, in his first paper on wave mechanics received about one month after Lanczos’s paper, a differential equation as the equation of motion of the de Broglie wave, and showed the existence of quantum numbers by solving the eigen value problem of the differential equation. Lanczos’s kernel $K(s, \sigma)$ of the integral equation should correspond, as Schrödinger pointed out in his footnote mentioned above, to

$\text{Lanczos’s } E(s, \sigma) \text{ is identical with Dirac’s } \delta\text{-function, } \delta(s - \sigma). \text{ Equation (162) due to Lanczos amounts to the continuous spectral representation of “Born-Jordan’s quantum condition”, namely the commutation relation (Cf. Sec. 2, §4, Chap. 3 of the present volume).}$
the Green function of the Schrödinger equation, that is, the inverse operator $H^{-1}$ of the Hamiltonian given as a differential operator.

§2. **Born-Wiener’s attempt by means of operators**

In their paper received on Jan. 5, 1926, "A new Formulation of the Quantum Laws for Periodic and Aperiodic Processes",\(^{77}\) Born and Wiener take up the problem that “the representation of the quantum laws by matrices meets with difficulties in aperiodic processes”. In the extreme case of a linear uniform motion, for example, “the coordinate matrix $q$ must have no element other than the diagonal, $m = n$, element, since there exists no period whatever, and this is impossible with continuous $m$ and $n$ in the exact sense”.

Though by matrix mechanics, in which the aggregate of the quantum amplitudes of a physical quantity, each having a frequency corresponding to a quantum transition, is grasped to be a matrix, the quantum condition in the old quantum mechanics is developed into a commutation relation, matrix mechanics remains to be the mechanics for periodic phenomena, just as the old quantum mechanics. We have described in Sec. 1, §4, Chap. 4, Vol. II the difficulty, which was pointed out by Einstein and Ehrenfest in 1922, of interpreting by means of the old quantum mechanics the result of Stern-Gerlach’s experiment. We have mentioned there that “the quantum condition did not furnish a dynamical law for the motion of translation of an unexcited or excited atom in a magnetic field”. The difficulty which Born and Wiener indicate in their paper under consideration is of the same kind as that which we have mentioned there.

Born-Wiener note first that the matrix $q$ can be seen as an “operation”, namely as a linear transformation of the set of variables $x_1, x_2, \ldots$

$$y_m = \sum_n q_{mn} x_n,$$

(163)

as is pointed out by Born-Heisenberg-Jordan in their theory of perturbation (cf. Sec. 2, §4 of the present chapter). Born-Wiener then consider the transformation of the following function $x(t)$ of time, since “the physical meaning is not clear” of the variables $x_1, x_2, \ldots$. They say the $x(t)$ “can be interpreted more easily”.

\(^{77}\)M. Born and N. Wiener, Zeit. f. Phys. 36 (1926), 174.
Attempts at the Interpretation of Matrix Mechanics

\[ x(t) \] is defined to be

\[ x(t) = \sum_n x_n e^{(2\pi i/\hbar)W_n t}. \] (164)

\( x_n \) is obtained inversely from \( x(t) \) by the averaging procedure

\[ x_n = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(s)e^{-(2\pi i/\hbar)W_n s} ds. \] (165)

Defining \( y(t) \) in the same way in terms of \( y_m \), one gets

\[
\begin{cases}
  y(t) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} q(t, s)x(s)ds, \\
  q(t, s) = \sum_{mn} q_{mn} e^{(2\pi i/\hbar)(W_m t - W_n s)},
\end{cases}
\] (166)

with the use of Eq. (163). The “operator”

\[ q = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} ds q(t, s) \ldots \] (167)

thus transforms the function \( x(t) \) into the function \( y(t) \). This is expressed in short form as

\[ y(t) = qx(t). \] (168)

\( q(t, s) \) acts as the generating function of the operator \( q \).

Denoting \( y(t) \) which is obtained from Eq. (166) with \( x(s) = e^{(2\pi i/\hbar)W_k s} \) by \( y_k(t) \) and putting \( q_k(t) = e^{-(2\pi i/\hbar)W_k t}y_k(t) \), one has

\[ q_k(t) = \sum_m q_{mk} e^{2\pi i\nu(mk)t}; \quad \hbar\nu(mk) = W_m - W_k. \] (169)

From \( q_k(t) \) the matrix element \( q_{nk} \) is obtained by means of

\[ a_{nk} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} q_k(t)e^{-2\pi i\nu(nk)t} dt. \] (170)

In order to generalize this relation, Born-Wiener express the function which results from the action of the operator \( e^{-(2\pi i/\hbar)W t}q \) on the function \( e^{(2\pi i/\hbar)W t} \) as

\[ q(t, W) = e^{-(2\pi i/\hbar)W t}qe^{(2\pi i/\hbar)W t}. \] (171)
From this $q(t, W)$ one gets the matrix "element"

$$q(V, W) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} q(t, W) e^{-\frac{(2\pi i/\hbar)(V-W)t}{T}} dt$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{-\frac{(2\pi i/\hbar)Vt}{T}} q e^{\frac{(2\pi i/\hbar)Wt}{T}} dt$$

(172)

according to Eq. (172). "The function $q(t, W)$ can exist in such cases where the matrix $q(V, W)$ does not exist in the true sense", and therefore "one can say that $q(t, W)$ represents a 'motion with the energy $W'"."

Because one has, with $D = d/dt$,

$$Dq x(t) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \frac{\partial q(t, s)}{\partial t} x(s) ds,$$

$Dq$ is an operator of the form given in Eq. (166), and its generating function is given by

$$\frac{\partial q(t, s)}{\partial t} = \frac{2\pi i}{\hbar} \sum_{mn} q_{mn} W_m e^{\frac{(2\pi i/\hbar)(W_m t - W_n s)}{T}} ds.$$

On the other hand, from the calculation of $z(x) = qDx(t)$ one gets for the generating function of $qD$

$$-\frac{\partial q(t, s)}{\partial s} = \frac{2\pi i}{\hbar} \sum_{mn} q_{mn} W_n e^{\frac{(2\pi i/\hbar)(W_m t - W_n s)}{T}}.$$

It is seen, therefore, that $Dq$ results from the matrix $(2\pi i/\hbar)Wq$, while $qD$ from the matrix $(2\pi i/\hbar)qW$, in terms of the diagonal matrix $W$, so that $D$ is relevant to the matrix $(2\pi i/\hbar)W$ and $Dq - qD$ to the matrix $\dot{q} = (2\pi i/\hbar)(Wq - qW).^{78}$

Born-Wiener thus suppose the time derivative of the operator $q$ to be

$$\dot{q} = Dq - qD.$$  

(173)

Let $q$ and $p$ be operators of the coordinate and momentum, respectively, satisfying the "commutation relation (quantum condition)"

$$pq - qp = \hbar/2\pi i.$$  

(174)

---

78) The elements of the matrix $(2\pi i/\hbar)(Wq - qW)$ are given by $2\pi i \nu(mn)q_{mn}$'s.
For an harmonic oscillator one has the equation of motion
\begin{equation}
\dot{q} + \omega_0^2 q = 0, \tag{175}
\end{equation}
whose solution is
\begin{equation}
q = e^{i\omega_0 t} \varphi_1(D) + e^{-i\omega_0 t} \varphi_2(D), \tag{176}
\end{equation}
where \( \varphi_1(D) \) and \( \varphi_2(D) \) are functions of \( D \).

This is because one has
\begin{equation}
\dot{q} = Dq - qD = i\omega_0 \{ e^{i\omega_0 t} \varphi_1(D) - e^{-i\omega_0 t} \varphi_2(D) \},
\end{equation}
from which Eq. (175) is obtained by means of the relation \( \dot{q} = D\dot{q} - \dot{q}D \).

Inserting Eq. (176) together with the relation \( p = \mu \dot{q} \) into Eq. (174), one gets
\begin{equation}
\frac{\hbar}{4\pi \mu \omega_0} = |\varphi_1(D)|^2 - |\varphi_1(D - i\omega_0)|^2, \tag{177}
\end{equation}
by making use of the relation \( \varphi_2(D + i\omega_0) = \varphi_1(D) \), which comes from the Hermitian property of the operator \( q \).\(^{79}\)

Operating Eq. (177) on the function \( e^{(2\pi i/\hbar)Wt} \), one gets the relation
\begin{equation}
\frac{\hbar}{4\pi \mu \omega_0} = \left| \varphi_1 \left( \frac{2\pi iW}{\hbar} \right) \right|^2 - \left| \varphi_1 \left( \frac{2\pi iW}{\hbar} - i\omega_0 \right) \right|^2. \tag{178}
\end{equation}
Because this difference equation shows that the difference in \( |\varphi_1|^2 \) is constant for every difference in \( W \) of \( h\omega_0/2\pi \), the value of \( W \) should be restricted to a certain domain for \( |\varphi_1|^2 \) not to be negative. If the value of \( W \) that makes \( |\varphi_1|^2 \) non-negative for the first time is taken to be zero, \( W \) takes the values \( W_k = kh\omega_0/2\pi \), where \( k = 0, 1, 2, \ldots \); (for \( \varphi_2 \), one has \( k = 1, 2, 3, \ldots \)). \( |\varphi_1(2\pi iW_k/\hbar)|^2 \) and \( |\varphi_2(2\pi iW_{k+1}/\hbar)|^2 \) are equally given by \( (\hbar/4\pi \mu \omega_0)(k+1) \).

By making use of these results together with Eq. (172), Born-Wiener derive the formula
\begin{equation}
q_{kl} \equiv q(W_k, W_l) = C \sqrt{k + 1} \delta_{k+1,l} + \sqrt{k} \delta_{k-1,l}, \tag{179}
\end{equation}
where \( C = (\hbar/4\pi \mu \omega_0)^{1/2} \), and show that the matrix relevant to \( H = p^2/2\mu + \mu^2 \omega_0^2 q^2/2 \) is diagonal with diagonal elements \( H_{kk} = (k+1/2)\omega_0/2\pi(k = 0, 1, 2, \ldots) \). They say that "Heisenberg's result is thus found again, and at the

\(^{79}\)Born-Wiener call the operator \( q \) Hermitian, when the matrix relevant to \( q \) is Hermitian.
same time it is shown that our calculation here provides of itself the discontinuous matrix as the unique solution”.

Having made these preparations, Born-Wiener now treat the “one-dimensional inertial motion” that is the main object of their study. In this case, the equation of motion becomes

\[ \ddot{q} = 0. \]

One can express its solution in the form

\[ q = t\varphi_1(D) + \varphi_2(D), \] (180)

as one has \( \dot{q} = Dq - qD = \varphi_1(D) \) and \( \ddot{q} = D\dot{q} - \dot{q}D = 0 \). From the Hermitian property of \( q \), it follows that \( \varphi_1(D) = \bar{\varphi}_1(D) \) and \( \varphi_2(D) = \bar{\varphi}_1(D) + \bar{\varphi}_2(D) \), where \( \varphi_1(x) \) denotes the derivative of the function \( \varphi_1(x) \). Thus, the “quantum condition” Eq. (174) becomes

\[ \frac{h}{2\pi i\mu} = \varphi'_1(D)\varphi_1(D), \] (181)

if the relation \( p = u\dot{q} \) is used. By operating Eq. (181) on the function \( e^{(2\pi i/h)Wt} \), one gets

\[ \frac{h}{2\pi i\mu} = \varphi'_1(u)\varphi_1(u); \quad u = \left(\frac{2\pi i}{h}\right)W. \] (182)

The solution to this differential equation is \( \varphi_1(u) = (hu/\pi i\mu)^{1/2} \), with which one has \( \varphi_2(u) = -\varphi'_1(u)/2 + \psi(u) \), where \( \psi(u) \) is a real function. These results together with Eq. (171) give

\[ q(t, W) = t\sqrt{2W/\mu} + \frac{ih}{4\pi \mu \sqrt{W}} + \varphi \left(\frac{2\pi iW}{h}\right). \] (183)

When \( W \) is large enough, the second term in the right-hand side of this equation can be neglected, so that the resultant \( q(t, W) \) and the formula

\[ p(t, W) = \sqrt{2\mu W} \] (184)

which derives from it “agree completely with the classical formula”.

In this way, Born-Wiener arrive at their conclusion that “this motion contains just no periodic component at all; nevertheless, it is accessible to our method for periodic motion”. Though \( q(t) \) may be more suitable for the image
based on classical theory than the aggregate of $q_{nm}$'s, as according to Born-Wiener's statement "our representation approaches classical theory so closely that it appears to be compatible with the bases of quantum theory", the physical meaning of $x(t)$ on which $q(t)$ is to act, or that of $x_n$ on which $x(t)$ is based remains unclarified.\textsuperscript{80)

However, for the groundwork for the attempt to formulate quantum mechanics of inertial motion in the paper under consideration, Born applies soon after the wave equation presented by Schrödinger to the problem of scattering of particles, making great contribution to the interpretation of the physical meaning of the wave function. This will be described in Sec. 5 of the next chapter. While Born-Wiener suppose that the operator $D = d/dt$ is relevant to the matrix $(2\pi i/h)W$, Schrödinger finds in his course of developing wave mechanics that the operators $-i(h/2\pi)\partial/\partial q$ and $+i(h/2\pi)\partial/\partial t$ acting on wave function correspond to $p$ and $W$ respectively. This process will also be described in the next chapter.

\textsuperscript{80) On this relation, confer Footnote 57) in Sec. 2, §7, Chap. 3 of the present volume.
Chapter 2

From the Proposal of Wave Mechanics to Quantum Mechanics

1. Schrödinger's Wave Equation

§1. Theory of gas based on de Broglie's wave theory

In the first half of the year 1926, Schrödinger proposed wave mechanics in a series of four papers under the title "Quantization as Eigenvalue Problem", and discussed the relation between the matrix element in matrix mechanics and the wave function in wave mechanics in a paper entitled "On the Relation of Heisenberg-Born-Jordan's Quantum Mechanics with Mine". In these papers, Schrödinger presented the equation of motion of the de Broglie wave, namely the differential equation which later came to be named the Schrödinger equation and taken to be the fundamental equation in quantum mechanics.

This work, however, was not Schrödinger's first work on the de Broglie wave. In his paper "On Einstein's Theory of Gas", 1) received on Dec. 15, 1925, about one month and a half before the first paper on wave mechanics, Schrödinger attempted to derive the results of Einstein's theory of gas based on Bose's method of statistical mechanics, 2) by treating gas molecules as standing de Broglie waves.

In this paper, Schrödinger says that "as the essential point in it new theory of gas, recently worked out by A. Einstein, it is quite generally regarded as that an entirely new statistics, the so-called Bose statistics, should be applied to the motion of gas molecules". Schrödinger considers that "this new statistics, as something primary, not to be seen as further explainable, resists one's

1) E. Schrödinger, Phys. Zeit. 27 (1926), 95.
2) Cf. Sec. 2, §4, Chap. 5, Vol. II.
natural feeling by right". He says that in the Bose statistics "a certain mutual dependence or an interaction of the molecules is covered".

This point was already pointed out by Einstein, as we have described in Chap. 5, Vol. II. Einstein thought that the fact that "the molecules are under mutual influence not understandable at present" was related to the wave nature of particles introduced by de Broglie. That is, Einstein attempted to understand the discovery of the new statistics, different from the classical statistics of particles, by means of the substantialistic idea of the de Broglie wave. From this viewpoint, Schrödinger treats now the motion of gas molecules as the eigen-oscillations of the de Broglie wave like those of the cavity radiation.

Schrödinger considers that "Einstein's theory of gas will be obtained by applying to the gas molecules the form of statistics that, when applied to 'light atoms', leads to Planck's law of radiation". Schrödinger says that "Planck's law of radiation can be gained, however, also through 'natural' statistics, by applying it to the so-called 'aether resonators', that is, to the degrees of freedom of the radiation", and cites Jeans's paper (1905)\(^3\) and Debye's paper (1910).\(^4\)

This means that, as we have seen in Sec. 1, §§2 and 3, Chap. 1, Vol. II, Jeans derived the Rayleigh-Jeans distribution by applying the equipartition law of energy to the eigen-oscillations of the cavity radiation, and Debye derived the Planck distribution by applying Planck's complexions to these eigen-oscillations, on the supposition that the energy would be partitioned in quanta of \(h\nu\) also for the cavity radiation as for Planck's resonators. The "natural statistics" in Schrödinger's statement means the method of partition function, as will be seen shortly below. If, for a system of oscillators, the partition function is calculated on the assumption that the energy of each oscillator takes only integer multiple values of \(h\nu\), Planck's mean energy of the system is obtained. In this procedure, the quantity \(h\nu\) is treated as the energy quantum, that is, as the "step" of the change of energy, and its connection with the light quantum is not taken into consideration.

Schrödinger says that in order to pass from this "natural statistics" into the Bose statistics, one has to "let the concepts of 'manifoldness of energy states' and 'manifoldness of the carriers of these states' exchange their roles". That is, he does not see \(h\nu\) simply as the "step" of the change of energy, but grasps it as the energy of a light quantum. "Then, the natural statistics — for example, Planck's method of summing over states — will lead to Einstein's

\(^3\)J. H. Jeans, Phil. Mag. 10 (1905), 91 [the same as Ref. 14] in Chap. 1, Vol. II.

\(^4\)P. Debye, Ann. d. Phys. 33 (1910), 1427 [the same as Refs. 31) and 50) in Chaps. 1 and 5, respectively, Vol. II.
theory of gas". Schrödinger states, however, that "this means only one thing, that de Broglie–Einstein’s wave theory of moving particle is serious, according to which a moving particle is nothing more than a kind of ‘foam crest’ on a wave radiation building the world ground".

A "foam crest on a wave radiation" means here a wave packet of the de Broglie wave. As has been mentioned in Chap. 5, Vol. II, the velocity of a wave packet is given by the group velocity, which is, in the case of the de Broglie wave, equal to that of a particle. Thus, Schrödinger considers a gas an aggregate of wave packets of the de Broglie wave, corresponding to the consideration of cavity radiation as an aggregate of light quanta. He makes analyses of the thermal property of a gas, with the use of a partition function (the sum over states), by relating the energies of the gas molecules needed for the calculation of the partition function to the group velocity.

Let us suppose that an ideal gas of monoatomic molecules is enclosed in a volume $V$. Every molecule can take states of energies $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s, \ldots$. It is assumed that "every state can have an arbitrary number of molecules at any time". If one considers that "the $s$th degree of freedom has the energy $n_s \varepsilon_s$", when there are just $n_s$ molecules in the state $\varepsilon_s$, "the $s$th degree of freedom behaves as an one-dimensional harmonic oscillator". "For the calculation, the whole system is therefore an aggregate of linear oscillators like a solid body, or to be precise, like a volume of radiation".

The partition function of the whole system is consequently given by

$$Z = \prod_s \sum_{n_s=0}^{\infty} e^{-n_s \varepsilon_s / kT} = \prod_s 1/(1 - e^{-\varepsilon_s / kT}),$$

(1)

with the restriction

$$\sum n_s = n,$$

(2)

where $n$ is the total number of molecules.\(^5\) With this restriction, the partition function $Z$ is expressed as

$$Z = \frac{1}{2\pi i} \int_{z=0} dz z^{-n-1} \prod_s \frac{1}{1 - z x_s},$$

(3)

\(^5\)With this restriction, $n_s$ is allowed to take the values 0, 1, 2, .... Because of this, Schrödinger obtained the same results as in Einstein’s theory of gas, namely the theory of gas based on the Bose-Einstein statistics. $n_s$ is restricted to take only the values 0 and 1 in the Fermi-Dirac statistics, which was discovered later than Schrödinger’s study under consideration (cf. Chap. 3 of the present volume).
when the residue theorem of complex integration is used, $x_s$ being defined as $x_s = e^{-\varepsilon_s/kT}$. If the method of saddle point is applied to evaluate the integral, it results that "Planck’s $\tilde{\Psi}$-function", that is $\tilde{\Psi} = k \log Z$, or $\tilde{\Psi} = -F/T$ in term of the free energy $F$, is given by

$$\tilde{\Psi} = -k \left[ \sum \log(1 - rx_s) + (n + 1) \log r \frac{1}{2} \log \left\{ 2\pi r \sum \frac{x_s}{(1 - rx_s)^2} \right\} \right], \quad (4)$$

in which $r$ is the value of $z$ that maximizes the integrand in Eq. (3) and is determined by the relation

$$(n + 1) = r \sum \frac{x_s}{(1 - rx_s)}. \quad (5)$$

One sees that Eq. (4) coincides with the relation

$$F = E - TS = kT \left\{ \sum \log(1 - e^{-\alpha s}) - An \right\}; \quad \alpha s = A + \varepsilon_s/kT, \quad (6)$$

which was derived in Einstein’s theory of gas, if one neglects the last term in Eq. (4) and puts $r = e^{-A}$. It is “completely of no importance” that one has $(n + 1)$ instead of $n$ for the coefficient of the second term in Eq. (4). The last term in Eq. (4) which is neglected is of the order of magnitude of $\log(n + 1)$. $r$ is identical with Einstein’s “parameter of degeneration” $\lambda = e^{-A}$.

Schrödinger then calculates $\varepsilon_s$ “in a close connection with de Broglie”. A molecule of the rest mass $m$ moving with a velocity $w = \beta c$ is “nothing more than a ‘signal’, or so to speak, ‘the foam crest’ of a wave system” whose frequency lies “in the neighborhood” of

$$\nu = \frac{mc^2}{\hbar \sqrt{1 - \beta^2}}. \quad (7)$$

For its phase velocity $u$, there holds the dispersion formula

$$u = \frac{c}{\beta} = \frac{c^2}{w}, \quad (8)$$

where “$w$ plays the role of the signal velocity”. Schrödinger numbers the $s$th energy state with “the number of the eigen oscillations” $s$ of the wave motion

---


7) Cf. Sec. 2, §2, Chap. 5, Vol. II. The signal velocity is another name for the group velocity, and derives from the fact that the group velocity determines the velocity with which the energy of a signal is transmitted.
in a volume \( V \), which is subject to this dispersion formula. The number of the
eigen oscillations with wavelengths between \( \infty \) and \( \lambda \) is given by

\[
s = \left( \frac{4\pi V}{3} \right) \lambda^3.
\]

(9)

With \( \lambda = u/v \), this gives

\[
s = \left( \frac{4\pi V}{3} \right) \frac{m^3 c^3 \beta^3}{h^3 \sqrt{1 - \beta^2}}.
\]

(10)

The energy quantum of the \( s \)th eigen oscillation is then given by

\[
\varepsilon_s = \hbar \nu_s = \frac{mc^2}{\sqrt{1 - \beta^2}} = mc^2 \left( 1 + \frac{J^2 s^{2/3}}{m^2 c^2} \right)^{1/2},
\]

(11)

in terms of \( J = \hbar (4\pi V/3)^{-1/3} \).

When \( \beta \) is small, one gets

\[
\varepsilon_s = (2m)^{-1} \hbar^2 \left( \frac{4\pi V}{3} \right)^{-2/3} s^{2/3},
\]

(12)

using only the kinetic energy for \( \varepsilon_s \). Equation (12) is identical with the expression given by Einstein. When \( \beta \) is close to unity, one has

\[
\hbar \nu_s = Jcs^{1/3}; \quad s = \left( \frac{4\pi V}{3} \right) \frac{\nu_s^3}{c^3}.
\]

(13)

"Apart from the factor of polarization 2", this is identical with the well-known
formula for the degree of freedom of the cavity radiation. Schrödinger says
that in Eq. (12) the value \( s = 0 \) is not allowed. This is because it corresponds
to \( \lambda = \infty \), while "the wavelength of the eigen oscillation with the longest
wavelength is the order of magnitude \( V^{1/3} \)."

The average value of \( n_s n'_s n''_s \cdots \) can be calculated with the use of the \( \Psi \)-
function:

\[
\langle n_s n'_s n''_s \cdots \rangle = \left( \frac{\partial}{\partial \log x_s} \right) \left( \frac{\partial}{\partial \log x'_s} \right) \left( \frac{\partial}{\partial \log x''_s} \right) \cdots \frac{\Psi}{k}.
\]

(14)

Neglecting the last term in Eq. (4) as before, one gets

\[
\frac{\{ \bar{n}_s^2 - (\bar{n}_s)^2 \}}{(\bar{n}_s)^2} = 1 + \frac{1}{\bar{n}_s},
\]

(15)
From the Proposal of Wave Mechanics to Quantum Mechanics

together with other expressions of similar type. They coincide with those obtained by Einstein.

dede Broglie treated the signal wave resulting from superposition of a large number of plane waves with a common wave normal of closely similar frequencies. It is a wave packet almost exclusively restricted to a thin domain between two parallel planes, but "the signal is to be restricted in a part of the space that is small in all the directions". In order to have such a wave packet, it suffices to let vary not only the frequency but also the wave normal in a small solid angle, as was done by Debye\(^8\) and Laue\(^9\) in the case of electromagnetic wave. Schrödinger says that "according to the classical laws of wave, it is certainly not achieved that the 'model of light quantum' so generated also remains always together". "It scatters rather into greater and greater space after passing a focus". However, he considers that "if one could avoid this last conclusion by a quantum-theoretical modification of the classical laws of wave, there would appear a way opened to the liberation from the dilemma of light quantum", expecting the effectiveness of the wave-theoretical view.

\section*{§2. The wave equation and quantum numbers}

Schrödinger presented in the first paper on "Quantization as Eigenvalue Problem",\(^10\) received on Jan. 27, 1926, the differential equation for standing waves of the de Broglie wave field. This equation later became known as the Schrödinger equation. By means of the eigenvalues of this differential equation, Schrödinger derived the Balmer formula of the energy levels of hydrogen atom.

In this paper, Schrödinger "shows that the usual quantization rule can be replaced by another demand, in which no word occurs of 'integers' anymore". As we have seen in Chap. 3, Vol. II, the existence of quantum numbers was premised in the quantum condition. It was Brillouin who first attempted in 1919, as we have described in Sec. 2, §1, Chap. 5, Vol. II, to introduce a kind of wave field into the atom in order to derive the quantum condition from the motion of a continuous body.

That is, Brillouin supposed an elastic wave propagated in the atom with a small velocity and tried to understand the quantum number \(n\) premised in the quantum condition, by considering the number \(n\) of times of the atomic

electron’s encounter with the wave which it emitted at a point on its orbit during the orbital motion until the wave passed through the atom.

de Broglie adopted Brillouin’s idea of understanding the quantum condition by the supposition of a wave field different from that of light, and assumed that only such orbits of the electron should be allowed that the phase wave propagating along respective orbits with a velocity greater than the light velocity would overtake the electron with one and the same phase at all times. For a nonrelativistic motion of the electron, this amounts to understanding the quantum condition as requiring the path length of every orbit should be an integer multiple of the wave length of the phase wave.

After de Broglie’s idea, Schrödinger sets up the wave equation and shows the existence of discrete energy levels as its eigen values. In Schrödinger’s theory, “the integer nature arises in the same way as the integer nature of the number of knots of an oscillating string, for example”. He starts his study by saying that according to his belief “the new view is capable of generalizing and touches very deeply the true essence of the quantum rules”.

“The usual form of the quantum rules is tied on to Hamilton’s partial differential equation”

\[ H\left(q, \frac{\partial S}{\partial q}\right) = E. \]  (16)

Usually, “of this equation a solution is looked for, which is represented as a sum of functions of a single one of the independent variables \( q \)”. However, Schrödinger “introduces now for \( S \) a new unknown \( \psi \) such that \( \psi \) would appear as a product of respective functions of the individual coordinates”. That is, he puts

\[ S = K \log \psi, \]  (17)

where \( K \) is a constant having the dimensions of action. Inserting this into Eq. (16) he gets

\[ H\left(q, \frac{K \partial \psi}{\psi \partial q}\right) = E. \]  (18)

\[ ^{11}\]It is needless to say \( q \) is the abbreviation of the Lagrange coordinates \( q_1, q_2, \ldots \), and \( H, E \) and \( S \) are respectively the Hamiltonian, energy and action function of the system. The notations \( S \) and \( W \) are used here by Schrödinger for the action function and Hamilton’s principal function, being inverse to our notations in Sec. 2, Chap. 3, Vol. II.
In the case of hydrogen atom, this equation becomes

\[ \left( \frac{\partial \psi}{\partial x} \right)^2 + \left( \frac{\partial \psi}{\partial y} \right)^2 + \left( \frac{\partial \psi}{\partial z} \right)^2 - \frac{2m}{K^2} \left( E + \frac{e^2}{r} \right) \psi^2 = 0, \]  

(19)

where \(-e\) and \(m\) are the electric charge and mass of the electron respectively and \(r = (x^2 + y^2 + z^2)^{1/2}\). For this differential equation in a modified form of classical physics, Schrödinger demands the variational condition

\[ \delta J = \delta \int \int \int dx dy dz \left\{ \left( \frac{\partial \psi}{\partial x} \right)^2 + \left( \frac{\partial \psi}{\partial y} \right)^2 + \left( \frac{\partial \psi}{\partial z} \right)^2 - \frac{2m}{K^2} \left( E + \frac{e^2}{r} \right) \psi^2 \right\} = 0, \]

(20)

and "with this variational problem replaces the quantum condition".\(^{12}\) With the use of partial integration, he gets

\[ \frac{1}{2} \delta J = \int df \delta \psi \frac{\partial \psi}{\partial N} \]

\[ - \int \int \int dx dy dz \delta \psi \left\{ \Delta \psi + \frac{2m}{K^2} \left( E + \frac{e^2}{r} \right) \psi \right\} = 0, \]

(21)

where \(\delta \psi\) is the variation of \(\psi\), \(df\) the surface element of a closed curved surface of infinite size, \(N\) the coordinate in the direction of its outward normal and \(\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2\). From this it results that

\[ \Delta \psi + \frac{2m}{K^2} \left( E + \frac{e^2}{r} \right) \psi = 0 \]

(22)

and

\[ \int df \delta \psi \frac{\partial \psi}{\partial N} = 0. \]

(23)

Generally speaking, Schrödinger derives the differential equation of \(\psi\) by "searching for such real, one-valued in the whole configuration space, twice-differentiable, finite and continuous functions \(\psi\) that make an integral over the whole configuration space of the so-called quadratic form an extremum".

\(^{12}\)He adds a footnote to say that he "does not miss that this formulation is not quite unambiguous". 
The differential equation (22) can be solved in spherical coordinates \((r, \theta, \phi)\) by setting \(\psi\) as a product of functions of single variables. “The method is fully known”. The dependence of \(\psi\) on the polar angles is given by spherical harmonics. For the function \(\chi(r)\) which represents the dependence on \(r\), one easily gets

\[
\frac{d^2 \chi}{dr^2} + \frac{2}{r} \frac{d \chi}{dr} + \left\{ \frac{2mE}{K^2} + \frac{2m^2}{K^2r} - \frac{n(n+1)}{r^2} \right\} \chi = 0, \tag{24}
\]

where \(n = 0, 1, 2, \ldots\). “The restriction of \(n\) being an integer is as is well-known necessary so the dependence on the polar angles becomes single-valued”.

Equation (24) has singular points at \(r = 0\) and \(r = \infty\), which constitute the boundary points of \(r\). Solutions which remain finite at both the boundary points can exist “only for certain outstanding values of the constants appearing in the equation”. Schrödinger makes his analysis by referring to Schlesinger’s book\(^\text{13}\) on differential equations. The method used is practically the same as those seen in present-day text books on quantum mechanics or mathematics for physics. As a result of the analysis Schrödinger finds that “for every positive \(E\) the Euler differential equation (22) of our variational problem has solutions, which are one-valued, finite and continuous in the whole space and go to zero as \(1/r\) in constant oscillations at the infinity”.\(^\text{14}\)

In the case of negative \(E\), there can be solutions only when \(E\) satisfies the condition

\[
\frac{me^2}{K\sqrt{-2mE}} = l; \quad l = 1, 2, 3 \ldots \tag{25}
\]

“The integer \(n\), which gives the order of the spherical harmonics, appearing in the solution should always be given only values smaller than \(l\)”. In the case of \(E < 0\), “the surface condition (23) will be guaranteed by its exponential vanishing at the infinity”. In the case of \(E > 0\), because \(\partial\psi/\partial N\) goes to zero only as \(1/r^2\) at the infinity, “the eigen functions provided do not solve the variational problem in the form proposed initially”. Schrödinger argues that “one must therefore still add a condition, for example, that \(\delta\psi\) vanishes at the infinity, or at least, that it should tend to a constant value irrespective of the direction so that the spherical harmonics bring the surface integral to zero”.

\(^\text{13}\) L. Schlesinger, “Differentialgleichungen” (Sammulung Schubert, Nr. 13, Göschen, 1903).
\(^\text{14}\) The number in the original paper of the equation (22) is (5).
From the condition (25), it results that

$$-E_l = \frac{me^4}{2K^2l^2}. \quad (26)$$

If one puts

$$K = \frac{\hbar}{2\pi}, \quad (27)$$

"this becomes the well-known Bohr energy level". Therefore, \( l \) is the principal quantum number. "\( n + 1 \) is analogous to the azimuthal quantum number, and the further splitting of this quantum number by the closer determination of the spherical harmonics can be considered analogous to the splitting of the azimuthal quantum number into an 'equatorial' quantum number and a 'polar' quantum number. These numbers determine here the system of the knot lines on the sphere". "Also the 'radial quantum number' \( l - n - 1 \) determines the number of "knot sphere"".

In this manner, Schrödinger arrives at the energy levels of the hydrogen atom, together with a verification of the existence of quantum numbers, from the conditions that \( \psi \) must satisfy. But, about the azimuthal and magnetic quantum numbers, their connection with the angular momentum is not yet understood with clarity. Nevertheless, it should be noted that Schrödinger showed, independently of Pauli who was based on matrix mechanics, that the "analogy" to the azimuthal quantum number is \( n + 1 \), differing by one from \( n \) itself. Equation (22) with \( K = \hbar/2\pi \) is nowadays called the time independent Schrödinger equation (for the hydrogen atom).

§3. The frequency of standing de Broglie wave and that of radiation

Schrödinger put the action function \( S \) in the form \( S = K \log \psi \) in order to obtain the wave equation for the de Broglie wave. This form is chosen because the Hamilton-Jacobi partial differential equation for the action function \( S \) is solved by separation of variables in the form of a sum \( S = \sum S_i(q_i) \), while the wave equation for a wave field \( \psi \) is solved by separation of variables in the form of a product \( \psi = \prod \psi_i(q_i) \). Schrödinger stresses this by writing the words "sum" and "product" in italics.

Schrödinger does not state any reason to "replace the quantum condition with the variational problem" given in Eq. (20), instead of using the modified
classical equation (19). However, we can infer his way of thinking by con-
sidering the case of eigen oscillations of a string. Let us take a string of a
uniform density $\rho$ stretched with tension $T$ between $x = a$ and $x = b$. For the
displacement $y(x, t)$ of the string, one has from the principle of least action,

$$
\delta \int_{t_1}^{t_2} \int_a^b \left\{ \frac{\rho}{2} \left( \frac{\partial y}{\partial t} \right)^2 - \frac{T}{2} \left( \frac{\partial y}{\partial x} \right)^2 \right\} \, dx \, dt = 0.
$$

(28)

When one treats only the standing waves, one puts $y(x, t) = \eta(x) \sin(2\pi vt + \alpha)$
and chooses the interval between $t_1$ and $t_2$ to be one period, getting

$$
\delta \int_a^b \left\{ \left( \frac{d\eta}{dx} \right)^2 - \left( \frac{2\pi}{u} \right)^2 \nu^2 \eta^2 \right\} \, dx = 0,
$$

(29)

where $u = \sqrt{T/\rho}$ is the wave velocity. By applying the partial integration,
one has

$$
\left[ \delta \eta \frac{d\eta}{dx} \right]^b_a - \int_a^b \delta \eta \left\{ \frac{d^2\eta}{dx^2} + \left( \frac{2\pi}{u} \right)^2 \nu^2 \eta \right\} \, dx = 0,
$$

(30)

where $\delta \eta$ denotes the variation of $\eta$. From this, there results the equation for
standing waves

$$
\frac{d^2\eta}{dx^2} + \left( \frac{2\pi}{u} \right)^2 \nu^2 \eta = 0,
$$

(31)

and the boundary condition

$$
\left[ \delta \eta \frac{d\eta}{dx} \right]^b_a = 0.
$$

(32)

Equations (29)–(32) given here are to be compared with Eqs. (20)–(23) due
to Schrödinger, respectively. This comparison shows that the factor $(E + e^2/r)$
in Eq. (20) corresponds in a certain sense to the eigen frequency squared.
It can be justified by the following arguments made by Schrödinger in the
third section of his paper under consideration, that our inference of his way of
thinking is on the mark.\textsuperscript{15}

In the third section, Schrödinger stresses first that “it suggests itself very
naturally that the function $\psi$ applies to a process of oscillation in the atom”.

\textsuperscript{15}In his second paper which we shall treat in the next section, Schrödinger mentions about
Eq. (20) that “the integrand of the stationary value integral is the Lagrangian of this process
of oscillation”.

He says that he “owes the stimulation for these considerations primarily to the brilliant thesis of Mr. Louis de Broglie”. “The main difference is that de Broglie thinks of progressive waves, while we are led to stationary eigen-oscillations when we put our formulae under the idea of oscillation”.

“If the individual solutions $\chi(r)$ of Eq. (24) multiplied by a spherical harmonics of order $n$ are grasped as the description of processes of oscillation, the quantity $E$ must have something to do with the frequency of the process in question”. “Now, it is usual in problems of oscillation that ‘the parameter’ (usually named $\lambda$) is proportional to the square of the frequency”. In the present case, however, firstly the frequency becomes imaginary for $E < 0$, and secondly “the feeling of a quantum-theorist tells him that the energy must be proportional to the frequency itself and not to its square”.

In order to “solve this contradiction”, Schrödinger considers that for the parameter $E$ in Eq. (22) “no natural zero level is fixed in advance, particularly because the unknown function $\psi$ appears to be multiplied apart from by $E$ still by a function of $r$, which may be changed by a constant with a corresponding change in the zero level of $E$”. “Consequently, the ‘expectation of the oscillation-theorists’ should be corrected so that not $E$ itself but $E$ added by a certain constant is expected to be proportional to the square of the frequency”.

Schrödinger supposes, thus, the frequency of the process of oscillation is given in the form

$$\nu = C'\sqrt{C+E} = C'\sqrt{C} + \left(\frac{C'}{2\sqrt{C}}\right)E + \cdots, \quad (33)$$

where $C$ is a constant which is large compared with $E$. He says that “this view has another merit very worthy of esteem”, and considers that “it mediates an understanding of Bohr’s frequency relation”. According to Bohr’s frequency relation, the frequency of radiation is proportional to the difference in $E$, and “therefore proportional to the difference between the eigen-frequencies $\nu$ of those hypothetical processes of oscillation”. “The frequency of radiation appears accordingly as a low ‘difference tone’ of eigen oscillations occurring with much higher frequencies”.

Schrödinger thinks it very understandable that “the light wave is causally linked with the beats appearing necessarily at every place of the space during the transition, and the frequency of light is determined by the number of times with which the maximum of the intensity of the beat returns per second”. He says that “it may arouse hesitation that… Bohr’s frequency relation is
apparently given the character of an approximate formula", but "this is only apparent and will be completely avoided when one develops the relativistic theory". In that case, "the large additive constant $C$ hangs naturally on the deepest connection with the rest energy $mc^2$ of the electron". He mentions also that it will be explained or avoided by the relativistic theory that the constant $h$, which is introduced into the theory through Eq. (27), should be separately introduced concerning Eq. (33) to give the frequency relation.

Schrödinger thinks that "the reality, which is doubted as regards electron orbits many times today, approaches in a greater degree the process of oscillation in the atom" to which the function $\psi$ is supposed to apply. de Broglie treated a wave that had a definite frequency proportional to the energy of an atom and was in progress with a certain phase velocity along a classical-theoretic orbit of the electron. On the other hand, Schrödinger assumes a standing wave that spreads throughout the atom in accordance with the wave equation proposed by him. However, the proportionality of its frequency to the energy of the electron is not given directly from the wave equation, but is deduced in an approximate way by analogy with classical oscillations in the form of Eq. (33), so that Schrödinger tries to connect the frequency of light to that of a beat between de Broglie waves.

According to this view, the atom oscillates in the form of an eigen-oscillation when it does not radiate. Schrödinger says that "this view however still differs very greatly from the natural image of an oscillating system, because a macroscopic system produces in general a pot-pourri of their eigen-oscillations". But he adds that "one may well say that only in the ground state the atom oscillates with an eigen-frequency and just for this reason does not radiate, because no beat appears". In his second paper, however, Schrödinger, from another viewpoint, again looks at Eq. (33) given for the relation between the frequency of a standing wave and the energy of the atomic electron, saying that "in the first communication this relation has arisen in the frame of mere speculation".

2. "Geometric Mechanics" and "Undulatory Mechanics"

§1. The Fermat principle and the Hamilton principle

Schrödinger attempts, in his second paper "Quantization as Eigenvalue Problem"\textsuperscript{16} received on Feb. 23, 1926, to "light up more closely the general

\textsuperscript{16}E. Schrödinger, Ann. d. Phys. 79 (1926), 489.
connection which exists between the Hamilton partial differential equation (H. P.) of a mechanical problem and the "belonged wave equation". He says that the demand for the variational principle taken in his first paper is "in itself nonunderstandable". As we have described in the last section, there arise certain difficulties from the supposition that the factor \( (E + e^2/r) \) multiplying the function \( \psi \) in the wave equation is connected with the frequency-squared of the standing wave in the atom.

Schrödinger starts his study in the second paper by pointing out "the inner connection of the Hamilton theory with the process of wave propagation". Citing Whittaker's book\(^{17}\) on analytical dynamics to argue that Hamilton developed his theory of mechanics from his theory of optics of inhomogeneous media, on the basis of this connection as the starting-point Schrödinger thinks that "the Hamilton variational principle can be grasped as the Fermat principle for wave propagation in the configuration space \((q\text{-space})\)". As we have described in Sec. 2, §3, Chap. 5, Vol. II, de Broglie noted this point in his discussions of the wave nature of the electron, and derived the quantum condition by introducing the phase wave traveling along the orbit of an atomic electron. He predicted also the diffraction of the electron, by reason of the possibility that there could be phenomena to which the Fermat principle would not apply. He did not, however, arrive at the setting up of the wave equation appropriate for the phase wave.

Schrödinger arrived, any way, at a wave equation of standing de Broglie waves, starting from the Hamilton-Jacobi differential equation. For the purpose of getting a penetrating understanding of the entirely new wave equation, he notes again the similarity between the Fermat principle and the Hamilton variation principle. At the same time, as he points out this similarity, Schrödinger mentions that "regretfully this powerful and fateful scope of Hamilton's idea has been robbed of its beautiful vivid robe as a superfluous accessory in most modern reproductions in favour of a mere colorless representation of the analytical relations". These words of Schrödinger are interesting because they show his substantialistic way of thinking.

Now, denoting Hamilton's principal function by \( W \), one has

\[
\frac{\partial W}{\partial t} + \mathcal{T} \left( q_k, \frac{\partial W}{\partial q_k} \right) + V(q_k) = 0. \tag{34}
\]

For a conservative system, denoting Hamilton's characteristic function by \( S \) as
\[
W = -Et + S(q_k),
\]
one gets from Eq. (34)
\[
2T \left( q_k, \frac{\partial W}{\partial q_k} \right) = 2(E - V). \tag{34'}
\]
In these equations, \( T(q_k, p_k) \) and \( V(q_k) \) are the kinetic and potential energies of the system, respectively, written as functions of the canonical coordinates and momenta \( q_k \) and \( p_k \) respectively, and \( E \) is the energy of the system. \( W \) in Eq. (34') can be replaced with \( S \), but Schrödinger "lets it stand so".

He denotes with \( \bar{T} \) the kinetic energy expressed as the function of \( q_k \) and the velocity \( \dot{q}_k \), instead of \( q_k \) and \( p_k \). By making use of a "non-Euclidean measure" for the \( q_k \)-space, Schrödinger puts its line element squared as
\[
ds^2 = \bar{T}(q_k, \dot{q}_k)dt^2. \tag{36}
\]
\( dt^2 \) in the right-hand side of this equation is external, so that this equation represents a quadratic form of \( dq_k \) by the relation \( dq_k = \dot{q}_k dt \). With the use of the covariant and contravariant vectors in this non-Euclidean measure, he defines the gradient operator \( \text{grad} \) and the divergence operator \( \text{div} \) and express Eq. (34') as
\[
(\text{grad} W)^2 = 2(E - V), \tag{37}
\]
or
\[
|\text{grad} W| = \sqrt{2(E - V)}. \tag{37'}
\]

Let us consider in the \( q \)-space a curved surface for which \( W = W_0 \). If, at each point on this surface, one "erects the normal and cuts out the stretch on it"
\[
ds = \frac{dW_0}{\sqrt{2(E - V))}}, \tag{38}
\]
\( ^{18} \)In the case of one-body problem, one has \( T = (1/2m)(p_x^2 + p_y^2 + p_z^2) \) with the use of Cartesian coordinates, getting Eq. (37') immediately. For a system consisting of particles of various masses, one has also Eq. (37') by putting \( q_x = \sqrt{m x} \), etc., if use is made of Cartesian coordinates. If one uses a system of curvilinear coordinates, one needs a certain mathematical formulation such as Schrödinger makes, because \( T \) depends on not only \( \dot{q}_k \) but also \( q_k \).
the terminal points of the normals meet the curved surface \( W_0 + dW_0 \). Because the change of \( W \) with time is given by Eq. (35), this means that the curved surface \( W = W_0 \) moves by the distance

\[
ds = \frac{Edt}{\sqrt{2(E - V)}},
\]

during the time \( dt \). That is, the equi-\( W \) surface has the normal velocity

\[
u = \frac{ds}{dt} = \frac{E}{\sqrt{2(E - V)}},
\]

From this, Schrödinger thinks that "our system of surfaces \( W = \text{const.} \) can be grasped as the system of wave fronts of a progressive but stationary wave motion in the \( q \)-space", "because the construction of the normals can obviously be replaced with a construction of Huygens' elementary waves and their envelope". It is well known from textbooks on general physics that Huygens explained the phenomena of refraction and reflection by considering that spherical waves spread from every point on a wave front of light to form a new wave front of light as their envelope. What Schrödinger means is that when waves starting from points of the curved surface \( W = W_0 \) spread to the respective radii given by Eq. (38), their envelope is the curved surface \( W = W_0 + dW_0 \).

Thus, "the action function \( W \) plays for our wave system the role of the phase". That is, Schrödinger sees the Hamilton-Jacobi partial differential equation as "the expression of the Huygens principle". Since the phase velocity is then given by Eq. (40), and the refraction index is proportional to its inverse, the Fermat principle is formulated in the present case as

\[
0 = \delta \int_{P_1}^{P_2} \frac{ds}{u} = \int_{P_1}^{P_2} \frac{\sqrt{2(E - V)}}{E} \, ds = \frac{1}{E} \delta \int_{t_1}^{t_2} 2T \, dt .
\]

This is nothing but the Hamilton principle in the Maupertuis form for a mechanical system of \( E = \text{const.} \). The "rays of light" that are the trajectories of the normals to the wave front form the "trajectories of the system" in the \( q \)-space, in agreement with the relation \( p_k = \partial W/\partial q_k \). This relation shows that a group of the trajectories of the system can be derived from a specified action function \( W \), "as a stream from its velocity potential".\(^{19}\)

In spite of these considerations, "an analogy of mechanics exists just with geometric optics, or, if one wants to be precise, with a very primitive wave

optics, and not with the fully extended wave optics”. Also, the Fermat principle itself is grasped “purely geometric-optically using solely the concept of refraction index”. The velocity of the representative point in the $q$-space of the system is given by

$$w = \sqrt{2T} = \sqrt{2(E-V)},$$

and is proportional to $1/u$ for $E = \text{const}$. The representative point cannot therefore remain on an equi-$W$ surface. “And so because even such important concepts in the wave theory as amplitude, wavelength and frequency — or, speaking more generally, the wave form — do not appear at all in the analogy, they lack a mechanical parallel”, says Schrödinger. “Of the wave function nothing is said at all, and for the waves $W$ has only the meaning — actually somewhat vague by the indetermination of the wave form — of the phase”.

Citing a paper of Sommerfeld and Runge, Schrödinger says that geometric optics is “a rough approximation” which applies to cases of short wavelength. Sommerfeld and Runge showed that if the amplitude of light $u$ is put in the form $u = u_0 \exp[2\pi i (\nu t + \alpha/n\lambda)]$ with the frequency $\nu$, wavelength $\lambda$, refraction index $n$, and position-dependent functions $u_0$ and $\alpha$, there will result the equation

$$\left( \frac{\partial \alpha}{\partial x} \right)^2 + \left( \frac{\partial \alpha}{\partial y} \right)^2 + \left( \frac{\partial \alpha}{\partial z} \right)^2 = n^2$$

from the wave equation, in the limit of short wavelengths. Schrödinger notes the fact that a quadratic first-order differential equation similar to the Hamilton-Jacobi partial differential equation is derived for the phase, as an approximation at short wavelengths to the wave equation which is a linear second-order partial differential equation.

Schrödinger states that, since “we know certainly today that our classical mechanics fails for orbits of very small dimensions and very strong curvatures”, as shown in his first paper, “it is essential to search for an undulatory mechanics”. Considering that “the immediate way to it is probably the wave-theoretical development of Hamilton's image”, he assumes the frequency of the wave in connection with $W$ as the phase, as will be described in the following subsection.

---

§2. Assumption of the wave frequency and the wave equation

"The extension of the q-space optics in the wave-theoretical sense" cannot be achieved by means only of the analogy between the Fermat principle and the Hamilton principle. Thus, Schrödinger makes the assumption that "it is a relevant extension of the analogy to regard the wave system considered above as sinusoidal waves". He adds that "this is the most simple and immediate, but must be arbitrary". On this assumption, "the wave function should contain the time only in the form of a factor sin(\cdots), whose argument is a linear function of W".

Because $W$ has the dimensions of action but the phase of a sinusoidal function is dimensionless, the coefficient of $W$ in this linear function must have the dimensions of the inverse of action. Schrödinger assumes that the coefficient is "universal, i.e. independent not only of $E$ but also of the nature of the mechanical system", and denotes it by $2\pi/h$. The time factor reads therefore

$$\sin\left(\frac{2\pi W}{h} + \text{const.}\right) = \sin\left\{\frac{2\pi Et}{h} + \frac{2\pi S(q_k)}{h} + \text{const.}\right\},$$

so that "the frequency of the q-space wave" is given by

$$\nu = \frac{E}{h}. \quad (45)$$

Its proportionality to the energy of the system thus results "without any noticeable affection" such as used in his first paper.

The wavelength then becomes

$$\lambda = \frac{u}{\nu} = \frac{h}{\sqrt{2(E-V)}}, \quad (46)$$

and is independent of a constant additive to energies. Schrödinger compares this wavelength $\lambda$ with the size $a$ of the electron orbit of the hydrogen atom. Because distances in the q-space have the dimensions of length multiplied by $\sqrt{\text{mass}}$, and $(E-V)$ is equal to the kinetic energy, the ratio of $\lambda$ to $a$ is given by

$$\frac{h}{m\sqrt{\text{mass}}} \quad (47)$$

\[^{21}]\text{Cf. Footnote 18) in the present chapter.}
where \( w \) denotes the velocity of the electron. The denominator \( mwa \) of this expression is of the same order of magnitude as the angular momentum of the electron, which amounts roughly to \( 10^{-24} \) cgs. according to the quantum theory. "We get therefore for the limit of the area of approximate validity actually the correct order of magnitude, if we identify our constant \( h \) with Planck's quantum of action". Equation (45) assumed by Schrödinger thus gives the same relation between frequency and energy as that of \( E = h\nu \) in the case of a light quantum.

Equation (45), together with Eq. (40), gives

\[
\frac{h\nu}{\sqrt{2(h\nu - V)}}.
\]

(48)

Because the velocity of the wave depends on the frequency, this equation gives "the dispersion law for the wave". "This dispersion law presents a great interest". This is because, according to this dispersion law, "the system velocity \( w \) now also has for the wave a very concrete meaning". Namely, \( w \) given in Eq. (42) is confirmed to obey the relation

\[
w = \frac{d\nu}{d(\nu/u)},
\]

(49)

which shows that "the velocity of a representative point of the system is that of a group of waves which fills a small region of frequency (signal velocity)".

In this way, "one again finds here a law, which Mr. de Broglie has derived for the 'phase wave' of the electron in his beautiful investigations". Schrödinger points out, however, that it has been derived by de Broglie "with essential reference to the theory of relativity", while here "it is a matter of a theory of greater generality which does not spring out of the theory of relativity, but also has the validity for every conservative system in the usual mechanics". This means that de Broglie arrived at his idea of phase wave, as we have seen in Chap. 5, Vol. II, by supposing an internal oscillation accompanying a particle which has in its rest system the frequency of \((\text{rest energy})/h\).

Of the fact that the velocity \( w \) of the representative point of the system is equal to the group velocity of the wave propagating with the velocity \( u \) in the \( q \)-space, Schrödinger discusses the case of a wave packet which at a given time passes through a given point in the \( q \)-space in a given direction. The mean frequency of the wave packet, or the mean energy of the system, is assumed to be given. For a mechanical system, this corresponds to the situation where
the particles belonging to the system start at a given time with respectively
given velocity components. The front of the $W$-wave is determined by the
same number of parameters ($E$ and others) as that of the degrees of freedom
of the system. He shows that if a wave packet is formed by superposing wave
fronts with the values of the parameters restricted in an infinitesimal domain,
the motion of the point at which the wave fronts are in phase obeys the same
law as that of the representative point of the system in the $q$-space.

Schrödinger states consequently that “the actual mechanical phenomenon
is in a relevant way grasped or reproduced by means of the wave phenomena in
the $q$-space”. But, he “thinks it a very difficult problem to prove exactly that
the superposition of this system of waves really gives a noticeable amplitude
only in a comparatively small neighbourhood of the point of coincident phase,
while it otherwise vanishes everywhere through interference”. Schrödinger later
solves this problem for a harmonic oscillator.

Even if “we have seen that for such a signal or a group of waves the same
equation of motion actually holds just as well as what classical mechanics
sets up for a representative point”, “in order to get down to an image of the
manifoldness of possible phenomena, a strict wave-theoretical treatment must
enter, namely one must start from the wave equation”. “The basic equations
in mechanics are useless for the explanation of microstructures of mechani­
cal phenomena, just as geometric optics is for the explanation of diffractive
phenomena”.

“The only datum to it set up” is the velocity of propagation $u$. “From this
datum the wave equation is naturally not tied down clearly”. So, Schrödinger
supposes that the wave equation must be a second order differential equation,
“only to induce endeavours according to simplicity”, and assumes for the wave
function the equation

$$\text{div}(\text{grad} \psi) - \frac{1}{u^2} \ddot{\psi} = 0,$$

(50)

though he mentions at the same time that “it is not fully agreed”. Because
“$\psi$ changes with time according to a factor $e^{2\pi ivt}$”, one gets, with the use of
Eqs. (45) and (48),

$$\text{div}(\text{grad} \psi) + \frac{8\pi^2}{\hbar^2} (E - V)\psi = 0.$$

(51)

The operators div and grad refer here, of course, to the configuration space.
"This is not the only one on the assumption of second order". It is possible to generalize \( \text{div}(\text{grad}\psi) \) to

\[
f(q_k) \text{div} \left\{ \frac{1}{f(q_k)} \text{grad}\psi \right\},
\]

where \( f(q_k) \) is an arbitrary function of \( q_k \). Schrödinger however uses Eq. (51) for the sake of "simplicity". He adds a footnote to say that "the introduction of \( f(q_k) \) means that not only the 'density' but also the 'elasticity' varies with the position".

As these words of Schrödinger show, he assumes Eq. (50) by analogy with the usual simple wave equation for elastic bodies. These circumstances are the same as those described of his first paper quoted in Sec. 1, §3 of the present chapter. The only difference is that, in his first paper a variational principle is assumed to derive an equation of standing waves similar to that of eigen oscillations of an elastic body, and the relation between the frequency and energy of the standing wave is also considered by analogy with an elastic body, whereas in his second paper the wave equation is set up by analogy with an elastic body through determining the relation between frequency and energy, on the assumption that the phase of the standing wave is proportional to Hamilton's principal function.

This analogy is, however, just an analogy so that the wave equation (51) is for "waves in the \( q \)-space". Schrödinger mentions that "one is not allowed to forget that it is only just about one-body problems that the interpretation as oscillations in the actual three-dimensional space is immediately permitted". He treats, as one of the "examples of application" given in the second paper under consideration, "the non-rigid rotator (diatomic molecule)", derives the wave equation for the motion of the center-of-mass and that for the relative motion by separation of variables with the use of the center-of-mass and relative coordinates, and shows their solutions. This makes the first case in which the Schrödinger equation is concretely applied to a two-body problem. Besides this example of application, Schrödinger discusses the examples of the Planck oscillator, the rotator with fixed axis, and the rigid rotator with free axis, illustrating that the wave equation (51) "contains the quantum condition", namely that the existence of quantum numbers is derived from the one-valued, finite and continuous properties of the wave function \( \psi \).
§3. Quantum condition, theory of Fresnel zones and WKB approximation

So Schrödinger leaves "geometric mechanics" for "undulatory mechanics", taking the analogy between the Hamilton principle and the Fermat principle as the footing, from the consideration that the equations of motion in classical mechanics "are useless for the explanation of the microstructure of mechanical phenomena". From this standpoint of Schrödinger, it seems "very characteristic" that the theory developed by Sommerfeld, Schwarzschild and Epstein is "in the most intimate connection" with the Hamilton-Jacobi partial differential equation and with the method of solving it. As we have seen in Chap. 3, Vol. II, the quantum condition was applied in their theory to the action integral $S$ obtained from the solution of the Hamilton-Jacobi partial differential equation.

Schrödinger says that in their theory "a certain interpretation of this microstructure in connection with classical mechanics, of course on very artificial additional assumptions, has in general succeeded and presented practical results of the highest significance". Moreover, he sees there that "the clearest hint is already contained of the true undulatory character of the mechanical phenomenon". Whereas the Hamilton-Jacobi partial differential equation corresponds to the Huygens principle, Schrödinger considers that "as the Huygens principle, completed by some rules quite nonunderstandable to geometric optics (the construction of Fresnel zones), becomes suitable for diffractive phenomena to a large extent, so from the theory of the action function lights could fall on the phenomena in atoms".

That is, Schrödinger regards the quantum condition as the analogy to the theory of Fresnel zones. In this theory, which is based on the wave theory, the effect of a wave front on a point is calculated by dividing the wave front into such narrow zones (Fresnel zones) that the distances from adjacent two of them to the point are different from each other by half a wavelength. Diffractive phenomena are explained by considering that the interference effect is fundamentally determined by whether the number of the Fresnel zones is even or odd.

As we have described in detail in Vol. II, the quantum condition was found after the process by which Nicholson arrived at the concept of the quantum nature of the angular momentum of an atomic electron rotating around the nucleus, and on the basis of this concept, Bohr grasped the existence of stationary
states of an atomic electron, and Sommerfeld interpreted the quantum nature of angular momentum as that of action integral to study general kinds of stationary states. Thus, originally the quantum condition had nothing to do with the wave theory. The reason why Schrödinger could regard the quantum condition as analogous to the theory of Fresnel zones is that, as has also been described in Vol. II, Brillouin tried to interpret the quantum condition by supposing a kind of elastic wave in an atom, and de Broglie adopted this idea to introduce the phase wave. In this sense, the analogy pointed out by Schrödinger is better said to be the analogy between the theory of Fresnel zones and the quantum condition under Brillouin-de Broglie’s interpretation.

The fact remains unchanged, however, that as the theory of Fresnel zones is “nonunderstandable” to geometric optics, so is the quantum condition to “geometric mechanics”. Schrödinger states that “one should get caught up in indissoluble contradictions when one tried to maintain the concept of the orbit of the system for these atomic phenomena, as much as one falls into nonunderstandability when one tries to follow the course of rays of light in the area of a diffractive phenomenon”, stressing the necessity of developing de Broglie’s wave theory.

The derivation of the quantum condition, which was regarded by Schrödinger as analogous to the theory of Fresnel zones, from Schrödinger’s wave equation was done by Wentzel in his paper “A Generalization of the Quantum Condition for the Purpose of the Wave Mechanics”, received on Jun. 18, 1926. In this paper, Wentzel treats at first a problem with one degree of freedom, and expresses Schrödinger’s wave equation as

\[ \psi'' + \left( \frac{2\pi}{h} \right)^2 p^2 \psi = 0; \quad p^2 = 2m\{E - V(x)\}. \] (52)

Putting

\[ \psi = e^{(2\pi i/h) \int y dx}, \] (53)

he gets

\[ \left( \frac{h}{2\pi i} \right) y' = p^2 - y^2. \] (54)

\[ ^{22)} \text{G. Wentzel, Zeit. f. Phys. 38 (1926), 518.} \]
This is a kind of the Riccati differential equation. Taking the limit of $h = 0$ and putting

$$\lim_{h \to 0} y = y_0 = \frac{dS}{dx},$$  \hspace{1cm} (55)$$

he derives $(dS/dx)^2 = y_0^2 = p^2$, which is the Hamilton-Jacobi differential equation in classical mechanics. Wentzel makes a power series expansion of $y$ in $h/2\pi i$:

$$y = \sum_{\nu=0}^{\infty} \left( \frac{h}{2\pi i} \right)^\nu y_{\nu},$$  \hspace{1cm} (56)$$
to get the recurrence formula

$$y'_{\nu-1} + \sum_{\alpha=0}^{\nu} y_\alpha y_{\nu-\alpha} = 0,$$  \hspace{1cm} (57)$$
which means

$$y_1 = -\frac{y_0'}{2y_0}, \hspace{1cm} y_2 = -\frac{y_1 + y_1^2}{2y_0^2}, \ldots .$$

In this way, two particular solutions of Eq. (52) are obtained starting from $y_0 = \pm p$.

He considers then the eigenfunction $\psi_k$ of the wave equation, which satisfies the boundary conditions. The value of the energy constant $E$ belonging to $\psi_k$ is the eigenvalue $E_k$. The boundary conditions demand that $\psi_k$ be finite or zero at the singular points of the differential equation. The domain of the variable $x$ is limited by the singular points. Because $y_0 = \pm p$ is pure imaginary outside the boundaries of the corresponding classical orbit, it turns out that when $x$ approaches a boundary point, one of the solutions obtained from $y_0 = \pm p$ by means of the recurrence formula becomes zero, and the other infinite.

The solution which satisfies the boundary condition at a singular point should be analytically continued to the other singular point, so as to satisfy the boundary condition there. Wentzel deduces the condition required for this purpose from "the theorem of oscillation", according to which an eigenfunction is characterized by the number of its knots (null points), and at every knot the function $\psi'_k/\psi_k$ has a pole of residue $2\pi i$. Taking a closed path of integration
enclosing all the knots, one has therefore

\[ \int y\,dx = kh, \quad (k = \text{integer} = \text{number of knots}). \]  

(58)

If one transforms the path so as to enclose also the singular points, one sees that for the particular solutions given above the result is of the same form as Eq. (58). In this way, Wentzel gets the following condition to be satisfied by an eigenfunction

\[ \sum_{\nu=0}^{\infty} \left( \frac{h}{2\pi i} \right)^\nu \int y_\nu\,dx = kh. \]  

(59)

In the limit of \( h = 0 \), this reduces to

\[ \int y_0\,dx = \int p\,dx = kh, \]  

(60)

in accordance with Sommerfeld’s quantum condition.

Wentzel then applies this method to the case of the hydrogen atom. Concerning the radial wave function, he shows that \( y_1 \) contributes \(-\hbar\) to “the sum of residues” in Eq. (59) at the singular point \( r = 0 \), but there is no contribution from \( y_2, y_3 \ldots \), so that Sommerfeld’s quantum condition remains unchanged. He also treats the Stark effect of the hydrogen atom and calculates \( y_0, y_1 \) and \( y_2 \) in the second order approximation for the electric field. His result is slightly different from the second order effect calculated by Epstein\(^{23}\) with the use of Sommerfeld’s quantum condition. Wentzel shows that his new result agrees better with the accurate experimental data obtained by Kiuti\(^{24}\).

In Wentzel’s study, it is not called to question that the singular points of the wave equation are in general different from the turning points of the particle in classical mechanics. In the case of a harmonic oscillator, for example, the singular points are \( x = \pm\infty \), whereas the turning points are \( x = \pm\sqrt{2E/K}, K \) being the elastic constant. Kramers studies this problem in his paper “Wavemechanics and half-integer Quantization”\(^{25}\) received on Sep. 9, 1926. Kramers, taking notice of the fact that, if \( p^2 \) is approximated by the linear function \( p^2 = \alpha_i (x - x_i) \) around the turning point \( x = x_i(i = 1, 2) \), \( y \) can be expressed by the


Bessel function of order 1/3, improves upon Wentzel's method to get
\[ \int p\,dx = \left( k + \frac{1}{2} \right) h, \]  
(61)

instead of Sommerfeld's quantum condition.

Brillouin takes, in his paper "The undulatory mechanics of Schrödinger; a general method of solution by successive approximations",\textsuperscript{26} received on Jul. 5, 1926, the general expression \( 2T = \sum m_{kl}q^kq^l = \sum m^{kl}p_kp_l \) for the kinetic energy \( T \) of the system, and writes the wave equation (51) in the configuration space in terms of \( m^{kl} \)'s, which are assumed to be dependent on the coordinates \( q_k \)'s. Putting \( \psi = \exp(2\pi iW/h) \), he expands \( W \) as a power series in \( h \) to get a general expression which corresponds to Wentzel's recurrence formula (57).

The method of approximate solution to the wave equation due to Wentzel, Kramers and Brillouin is nowadays called the WKB method. The WKB method confirmed Schrödinger's idea that the quantum condition should bear analogy with the theory of Fresnel zones in wave mechanics. It proved also that \( W \) defined by the relation \( \psi = \exp(2\pi iW/h) \) obeys, in the limit of \( h \to 0 \), the Hamilton-Jacobi partial differential equation, and therefore becomes identical with Hamilton's principal function. Thus, "the undulatory mechanics" characterized by Schrödinger's wave equation was shown to tend to "geometric mechanics", that is, to classical mechanics as \( h \to 0 \), and thereby Schrödinger's wave equation advanced in its epistemological significance.

Schrödinger regarded the Hamilton-Jacobi partial differential equation as the representation of the Huygens principle, on the basis of the similarity between the Hamilton principle and the Fermat principle. However, with this consideration alone "the wave equation was not tied down clearly". It was because of the demand by the experimental facts that Schrödinger was led to "identify the constant \( h \) with the Planck constant" in the relation assumed for the frequency \( \nu = E/h \), Eq. (45), and it was a jump on the ground of the analogy to elastic waves that he put the wave equation, Eq. (50), in the configuration space. We see, here, the dialectics of continuity \textit{versus} jump in the process of recognition of nature.

\textsuperscript{26}L. Brillouin, Comp. Rend. 183 (1926), 24.
3. Relation between Wave Mechanics and Matrix Mechanics

§1. Wave functions and matrix elements

Schrödinger made efforts to see matrix mechanics from the standpoint of wave mechanics in his paper "On the Relationship of Heisenberg-Born-Jordan's Quantum Mechanics to Mine" received on Mar. 18, 1926, about three weeks later than his second paper on wave mechanics. The works referred to in this paper regarding matrix mechanics are those of Heisenberg, Born-Jordan and Born-Heisenberg-Jordan, and the first and second papers of Dirac, all of which have been treated in the last chapter.

In this paper on the relationship, Schrödinger says that "Heisenberg's quantum mechanics" and "the undulatory or physical mechanics", both of these new quantum theories, "coincide with each other" with regard to special results known hitherto, even in the deviation of these results from those in the old quantum mechanics. It is meant here to indicate the appearance of the half-integer quantum number \((n + 1/2)\) in the case of a harmonic oscillator and of a rotator. He thinks that "this is very strange, because the starting points, ideas and whole mathematical apparatus seem actually to be completely different".

Schrödinger continues to discuss that "above all, however, the departure from classical mechanics in both theories seems to occur really in diametrically opposed directions". "In Heisenberg the classical continuous variables are replaced by systems of discrete numerical quantities (matrices), which, each being dependent on a pair of integer indexes, are determined by algebraic equations". In wave mechanics, on the other hand, "in place of the event describable by means of a finite number of total differential equations with a finite number of dependent variables, a continuous field-like event in the configuration space enters, which is controlled by a single partial differential equation deducible from a principle of action".

Schrödinger concludes, accordingly, that "the very intimate relationship between Heisenberg's quantum mechanics and wave mechanics should now be revealed". "From the formal mathematical standpoint one has to indicate it well as the identity (of both theories)". In order to prove this, he analyzes first "how one can assign a matrix to every function of the position and

---

27) E. Schrödinger, Ann. d. Phys. 79 (1926), 37. This paper will be called the "paper on the relationship" in the following.

28) The name of physical mechanics comes perhaps from that of physical optics for wave optics.
momentum coordinates, in such a way that these matrices satisfy in any case Born-Heisenberg’s formal rule of calculation”.

Schrödinger notes that “Heisenberg’s particular laws of calculation” for functions of the canonical coordinates and momenta \( q_1, q_2, \ldots q_n; p_1, p_2, \ldots p_n \) coincide precisely with “the laws of calculation which are valid according to the ordinary analysis for the differential operators” of the variables \( q_1, q_2, \ldots q_n \), and regards this as the starting point. In fact, the operator \( \partial / \partial q_l \) is “exchangeable” with \( \partial / \partial q_m \) for arbitrary \( m \), but with \( q_m \) only for \( m \neq l \). For \( m = l \) the operator \( (\partial / \partial q_l)q_l - q_l(\partial / \partial q_l) \) is identical with 1. “This simple fact is reproduced in the field of matrices as Heisenberg’s law of exchange”.

Now, Schrödinger uses the abbreviation \( F(q_k, p_k) \) for a function of \( q_k \)'s and \( p_k \)'s \((k = 1, 2, \ldots n)\), and supposes that it has a structure such as

\[
F(q_k, p_k) = f(q_k)p_rp_sp_tg(q_k)p_r'h(q_k)p_r''p_s'' \cdots \tag{62}
\]

He considers this as a “well-ordered functional symbol”, and assigns to it the following operator

\[
[F, \cdot] = f(q_k)K^3 \frac{\partial^3}{\partial q_r \partial q_s \partial q_t} g(q_k)K^2 \frac{\partial^2}{\partial q_r'' \partial q_s''} \cdots \tag{63}
\]

which is defined by replacing \( p_r \), etc. in Eq. (62) with \( K\partial / \partial q_r \), etc. \( K \) is a “universal constant”, and will later be put to be \( K = h/2\pi i \). The result of operating \([F, \cdot]\) on a function \( u(q_k) \) is expressed as \([F, u]\). If \( G \) is another well-ordered functional symbol, \([FG, u]\) is not in general equal to \([GF, u]\).

Schrödinger then takes a complete orthonormal set of an infinite number of functions in the whole \( q \)-space,

\[
u_1(x)\sqrt{\rho(x)}, \quad u_2(x)\sqrt{\rho(x)}, \quad u_3(x)\sqrt{\rho(x)}, \ldots \tag{64}
\]

where \( x \) shortly represents the \( q_k \)'s, and \( \rho(x) \) is a density function such as

\[
\int \rho(x)u_j(x)u_k(x)dx = \begin{cases} 0 & (j \neq k) \\ 1 & (j = k) \end{cases} \tag{65}
\]
with \( dx = dq_1 dq_2 \cdots dq_n \). Making use of this set of functions he assigns to the function \( F \) a matrix whose element is given by

\[
F^{kl} = \int \rho(x) u_k(x) [F, u_l(x)] dx . \tag{66}
\]

The indexes of a matrix are written here as superscripts to avoid confusion which may arise when matrices of such quantities as \( q_k, p_k \) and others are treated. Schrödinger defines the "rolled over operator" of \([F, \cdot] \) as

\[
[F, \cdot] = (-1)^{\tau} K^2 \frac{\partial^2}{\partial q_{\tau''} \partial q_{\tau'}} h(q_k) K \frac{\partial}{\partial q_r} g(q_k) K^3 \frac{\partial^3}{\partial q_r \partial q_s \partial q_l} f(q_k) , \tag{67}
\]

where \( \tau \) is the number of differentiations in \([F, \cdot] \). With the use of this, it can be shown by means of partial integration that \( F^{kl} \) can also be expressed as

\[
F^{kl} = \int u_l(x) \left[ \bar{\Phi}, \rho(x) u_k(x) \right] dx . \tag{68}
\]

Equation (68) gives

\[
\sum_l F^{kl} G^{lm} = \sum_l \int u_l(x) \left[ \bar{\Phi}, \rho(x) u_k(x) \right] dx \int \rho(x) u_l(x) [G, u_m(x)] dx ,
\]

which is transformed into

\[
\sum_l F^{kl} G^{lm} = \int \left[ \bar{\Phi}, \rho(x) u_k(x) \right] [G, u_m(x)] dx \tag{69}
\]

on account of the so called "completeness relation"\(^{29}\) for an orthonormal set of functions. The partial integration of the right-hand side of Eq. (69) results in

\[
\sum_l F^{kl} G^{lm} = \int \rho(x) u_k(x) [FG, u_m(x)] dx , \tag{70}
\]

which is equal to \((FG)^{km}\). Namely, the matrix defined by Eq. (66) satisfies the rule of product of matrices.

The operator assigned to the function \((p_l q_l - q_l p_l)\) is equal to \( K \) from Eq. (63), and one has

\[
(p_l q_l - q_l p_l)^{jk} = K \int \rho(x) u_j(x) u_k(x) dx = \begin{cases} 0 & (j \neq k) \[ K & (j = k) \end{cases} , \tag{71}
\]

\(^{29}\)R. Courant und D. Hilbert, "Methoden der matematischen Physik I" (Springer, 1924), S. 36.
which, if one puts \( K = \frac{\hbar}{2\pi i} \), is nothing but “Heisenberg’s quantum relation”. This result can be obtained, of course, also from

\[
\begin{align*}
q_{km}^l &= \int q_l(x) u_k(x) u_m(x) \, dx, \\
p_{km}^l &= K \int \rho(x) u_k(x) \frac{\partial u_m}{\partial q_l} \, dx,
\end{align*}
\]

with the use of the rules of product and sum of matrices. From a comparison of the operator \([\partial F/\partial q_l, \cdot]\) assigned to the derivative \(\partial F/\partial q_l\) of the function \(F\) with the operators \([p_l F, \cdot]\) and \([F p_l, \cdot]\), one sees that

\[
\left[ \frac{\partial F}{\partial q_l}, \cdot \right] = \frac{1}{K} [p_l F - F p_l, \cdot].
\]

One can also show for the operator assigned to \(\partial F/\partial p_l\) the relation

\[
\left[ \frac{\partial F}{\partial p_l}, \cdot \right] = \frac{1}{K} [F q_l - q_l F, \cdot],
\]

with the use of \((\partial/\partial q_l) q_l - q_l (\partial/\partial q_l) = 1\).

In this way, “matrices are constructed in a very general way”. Schrödinger shows, moreover, that “if one confers the role of mediation on a particular orthogonal set, namely the set of eigenfunctions of that partial differential equation which constitutes the basis of the undulatory mechanics, the system of the characteristic algebraic equations, which the authors of the matrix mechanics call ‘equations of motion’, are completely produced”, as follows.

§2. Mathematical equivalence between matrix mechanics and wave mechanics

Schrödinger takes a mechanical system characterized by a Hamiltonian \(H(q_k, p_k)\). In matrix mechanics, the matrices of \(q_k\) and \(p_k\) are supposed to obey the “equations of motion”

\[
\begin{align*}
\left( \frac{dq_l}{dt} \right)^j_k &= \left( \frac{\partial H}{\partial p_l} \right)^j_k, \quad (l = 1, 2, 3, \ldots n), \\
\left( \frac{dp_l}{dt} \right)^j_k &= \left( \frac{\partial H}{\partial q_l} \right)^j_k, \quad (j = 1, 2, 3, \ldots \infty).
\end{align*}
\]
The meaning of the right-hand side of these equations is "already explained", but "d/dt appearing in the left-hand side is not". To it matrix mechanics "attaches the meaning" that

\[
\left(\frac{dq_j}{dt}\right)^{jk} = 2\pi i (\nu_j - \nu_k) q_j^{jk}, \quad \left(\frac{dp_l}{dt}\right)^{jk} = 2\pi i (\nu_j - \nu_k) p_l^{jk},
\]

(76)

where \(\nu_j\) denotes the frequency of the state with the quantum number \(j\). According to Eq. (76) together with Eqs. (73) and (74), Eq. (76) becomes

\[
\begin{align*}
(\nu_j - \nu_k) q_j^{jk} &= (H q_l - q_l H)^{jk}/\hbar, \\
(\nu_j - \nu_k) p_l^{jk} &= (H p_l - p_l H)^{jk}/\hbar.
\end{align*}
\]

(75)

Now, if "the eigenfunctions of the natural boundary value problem" of the partial differential equation

\[
-[H, \psi] + E\psi = 0
\]

(77)

for the function \(\psi\) of \(\textbf{q}\)'s are chosen as the orthogonal set, it is seen that Eq. (75') is generally satisfied. For, if \(U_j\) and \(E_j\) denote the eigenfunction and eigenvalue of Eq. (77) respectively, one has

\[
[H, U_j] = E_j U_j,
\]

(78)

from which one gets, with the use of Eqs. (65) and (66),

\[
H^{kl} = E_k \int \rho(x) u_k(x) u_l(x) dx = \begin{cases} 0 & (l \neq k) \\ E_l & (l = k) \end{cases}
\]

(79)

and consequently

\[
\begin{align*}
(H q_l)^{jk} &= \sum_m H^{jm} q_m^{lk} = E_j q_l^{jk}, \\
(q_l H)^{jk} &= \sum_m q_l^{jm} H^{mk} = E_k q_l^{jk},
\end{align*}
\]

(80)

\[30\] If, in a Hamiltonian, there are terms such as \(q_k p_k^2\), for example, they are not "well-ordered". In matrix mechanics, they are symmetrized into such forms as \((q_k p_k^2 + p_k q_k p_k + p_k^2 q_k)/3\) with the use of the commutation relation, which are "well-ordered" functions. Schrödinger mentions that he "will not here enter to the general rule of symmetrization".
together with similar expressions for $p_i$. The frequency $\nu_j$ is given by the relation $\nu_j = E_j/h$.

$[H, \cdot]$ in Eq. (77) is the operator assigned to the (suitably symmetrized) Hamiltonian, but it is still necessary to show that $- [H, \cdot]$ coincides with the "wave operator" in wave mechanics. Schrödinger proves this by showing that if, in the Hamiltonian of the form

$$H = T(q_k, p_k) + V(q_k), \tag{81}$$

$T$ is given as a quadratic form of $q_k$'s, the wave equation takes the same form as Eq. (77), as follows.

The wave equation can be obtained from the variational method with subsidiary condition

$$\delta J_1 = \delta \int \left\{ \frac{\hbar^2}{4\pi^2} T \left( q_k, \frac{\partial \psi}{\partial q_k} \right) + \psi^2 V(q_k) \right\} \Delta_p^{-1/2} dx = 0, \tag{82}$$

$$J_2 = \int \psi^2 \Delta_p^{-1/2} dx = 1,$$

where $\Delta_p$ is the discriminant of the quadratic form $T$. Carrying out the variation with the use of a Lagrange multiplier $E$, one gets

$$0 = (\delta J_1 - E \delta J_2)/2 = \int \left\{ -\frac{\hbar^2}{8\pi^2} \sum_l \frac{\partial}{\partial q_l} \left[ \Delta_p^{-1/2} T_{pi} \left( q_k, \frac{\partial \psi}{\partial q_k} \right) \right] \right\} \delta \psi dx,$$

and hence the Euler equation

$$\frac{\hbar^2}{8\pi^2} \Delta_p^{1/2} \sum_l \frac{\partial}{\partial q_l} \left[ \Delta_p^{-1/2} T_{pi} \left( q_k, \frac{\partial \psi}{\partial q_k} \right) \right] - V(q_k) \psi + E \psi = 0, \tag{84}$$

where $T_{pi}$ is defined through the application of the Euler expression for homogeneous functions to the quadratic form $T$ by

$$T(q_k, p_k) = \sum_l p_l T_{pi}(q_k, p_k). \tag{85}$$

Equation (84) is the wave equation. The operator other than $E$ in the left side of this equation becomes, when $(\hbar/2\pi i)\partial/\partial q_k$ is replaced by $p_k$, identical with $[-H, \cdot]$ on account of Eqs. (81) and (85). Thus, "the solution of the whole
system of Heisenberg-Born-Jordan's matrix equations is led back to the natural boundary value problem of a linear partial differential equation. "If one solves the boundary value problem, one can work out every matrix element, in which one is interested, by differentiation and integration according to the instruction of Eq. (66)".  

About "the natural boundary value problem, that is, the natural boundary conditions at the natural boundary of the configuration space", Schrödinger makes reference to the examples worked out in his first and second papers.

He mentions that, in the paper on the relationship under consideration, he has restricted himself to the case of "classical mechanics without magnetic field", because to him "the relativistic magnetic generalization seems to be not yet sufficiently cleared up". He thinks, however, that "there is hardly any doubt about that the perfect parallelism of both the new quantum theories will remain to exist also for it".

Also in this paper on the relationship, the set of orthogonal functions is supposed to be "a set of perfectly discrete functions". Schrödinger says that "in actual cases of application it is certainly not so". Not only for the hydrogen atom but also for heavier atoms the wave equation must possess, besides a line spectrum, also "a continuous eigenvalue spectrum, which manifests itself among other things in the continuous optic spectra following the limit of spectral series". He mentions, however, that "the main aim of this note is to bring out the formal connection between the two theories clearly as far as possible, and this is surely not essentially altered by the appearance of a continuous spectrum". He adds "a caution that we should always keep in mind to assume the convergence of the expansion by the eigenfunctions".

§3. Phenomenology and substantialism

Having shown the mathematical equivalence of wave mechanics and matrix mechanics, Schrödinger discusses the "comparison of the two theories". "If both the theories should prove to be tenable in their present marked form, that is, to be already the correct generalization also for complicated systems, every discussion on the advantage of one or the other in certain sense has only an apparent subject. For, from the mathematical standpoint they are completely equivalent and it is a matter of fundamentally secondary question of the comfort of calculation".

31) The original number of Eq. (66) is (6).
Next, he states that “there are not a few physicists who see, quite in the sense of Kirchhoff and Mach, the task of the physical theory in a mathematical description, economical as far as possible, of the empirical connections among observable quantities, namely, a description which reproduces the connections without mediation of in principle unobservable elements as far as possible”. This indicates the idealistic epistemology in Mach's sense, the influence of which scientists at that time were under in no small way. According to the epistemology in Mach's sense, the cognition of nature consists in the process of getting mathematical laws by means of the arrangement of sensuous experiences. For this reason Schrödinger says also that “in such an attitude mathematical equivalence is almost synonymous with physical equivalence”.

From the standpoint in Mach's sense, thus, there is little to choose between wave mechanics and matrix mechanics. If any, as Schrödinger continues to say, “in the present case a certain advantage of the matrix representation could at most be seen in that, because of its complete nonintuitionality, it does not lead to forming space-time images of the atomistic events, which must perhaps remain uncontrollable in principle”. We remember that Heisenberg started his paper, which brought forth quantum mechanics, with the following words, as has been described in Sec. 1, §1 of the last chapter — “As is well known, against the formal rules which are generally used in the quantum theory for the calculation of observable quantities (the energy in hydrogen atom, for example), a serious objection is raised that these rules of calculation contain as essential constituents relations among quantities which cannot in principle be observed (such as the position and period of revolution of the electron, for example)”.

In spite of their mathematical equivalence, matrix elements are directly introduced in matrix mechanics as “quantities to substitute for classical quantities”, while they are shown in Schrödinger’s theory to be reproduced through the eigenfunctions. Regarding a “mathematical description economical as far as possible”, Heisenberg’s theory seems in the epistemology of Mach’s type to predominate over Schrödinger’s theory, because the concept of wave is introduced for atomic phenomena in the latter theory. However, “the equivalence exists also in the reverse direction”, as Schrödinger now goes to show.

He takes the problem of obtaining the function $u_j(x)$ in the equation

$$q_i^k = \int q_i u_j(k) u_k(x) dx,$$  \hspace{1cm} (86)
assuming that the quantity in the left side is known. [In this equation the density function \( \rho(x) \) is abbreviated when compared with Eq. (72).] According to the multiplication law of matrices, the quantity

\[
\int P(x)u_j(x)u_k(x)dx
\]

(87)
can be calculated for an arbitrary power \( P(x) \) of \( q_l \) with the use of the partial integration. The whole of these integrals for fixed \( j \) and \( k \) constitutes the "moments" of the function \( u_j(x)u_k(x) \). "One knows that on quite general assumptions a function is unambiguously fixed by the whole of its moments".

Consequently, \( u_j(x)u_k(x) \) is unambiguously fixed, inclusively of \( u_j^2(x) \) and therefore of \( u_j(x) \) itself.\(^{32}\) "Not only the matrices are constructed, as shown above, from the eigenfunctions, but also the eigenfunctions, reversely, from the numerically given matrices". The eigenfunctions are thus shown to be given by mediation of the matrices. Schrödinger states that "the wave functions do not form, for example, an arbitrary and special 'carnal dressing' of the bare skeleton of matrices, devoting to the need for the intuitionality".

In this way it turns out that, from the standpoint of the epistemology in Mach's sense, there is little to choose between wave mechanics and matrix mechanics. However, Schrödinger still thinks that "the thesis can be granted only conditional validity that mathematical equivalence should be synonymous with physical equivalence". He takes an example of two expressions for the electrostatic energy of a system of charged conductors. One of them is the space integral \( (1/2) \int E^2 d\tau \), and the other is the sum over the conductors \( (1/2) \sum e_i V_i \). These are perfectly equivalent for the electrostatics, the latter being obtained by means of a partial integration of the former. "Nevertheless we prefer the former as significant". This is because, whereas on the basis of electrostatics there is no reason to prefer one of them to the other, "in electrodynamics the first expression remains also useful, but the second does not". Namely, Schrödinger means the importance of the motion of electromagnetic field itself in electrodynamics.

Though Schrödinger thinks that "at present it can hardly be decided yet which of the two deserves preference from this viewpoint", he continues to analyze "the problems which, besides the original optical questions, are worth considering for further extension of the atomic dynamics". "They are placed

\(^{32}\)Schrödinger adds that, though it is necessary to remove the density \( \rho(x) \), "one will in any case have nothing to fear epistemologically".
to us about the experimental physics in eminently intuitional form, as for example: how do two colliding atoms or molecules rebound from each other, how is an electron or an α-particle deflected when, with a given velocity and an aerial velocity (‘perpendicular from the atom to the initial orbit’), it is shot through an atom?”

“In order to step closer to such problems, it is absolutely necessary to survey clearly the continuous transition between the macroscopic intuitional mechanics and the micromechanics of the atom”. Saying so, Schrödinger cited his second paper, in which he regards, as we have described in the last section, the Hamilton-Jacobi differential equation as the representation of the Heygens principle, and emphasizes the necessity to transit from geometric mechanics to wave mechanics, in accordance with the relative magnitude between the size of the objective system and the wavelength.

In the second paper he says also, about the assertion that the concept of the position and orbit of an electron should be removed and atomic events could not be incorporated into the space-time form of thinking, that he “would consider from the philosophical standpoint a conclusive decision in this sense to be equal to a complete laying down of arms”. He gives a simile that “a scholar of geometric optics who, in his continual study, approaching to diffractive phenomena by means of the concept of rays proven in the macroscopic optics, constantly fails, might at last perhaps come to the thinking that the geometric laws are not applicable to diffractive phenomena”.

“Light rays, known to the scholar as straight and independent of one another, now in a homogeneous medium suddenly show very strange bends and are evidently influenced mutually”. Schrödinger mentions “I think this analogy is very close”. “Indeed the reasonless bends are not missing the analogy in the atom — one thinks of the ‘nonmechanical compulsion’ brought out for the interpretation of the anomalous Zeeman effect”. Schrödinger’s words indicate the fact that, as we have described in Sec. 2, §3 and Sec. 3, §2, Vol. II, this idea proposed in 1923 by Bohr failed completely on account of the discovery of the electron spin.

Now, in the paper on the relationship between wave mechanics and matrix mechanics, Schrödinger continues to argue that to him “it seems extraordinary difficult to tackle problems of the kind described above,” so long as one feels obliged, on epistemological ground, to suppress the intuition in the atomic

33) Schrödinger means here atomic or molecular collisions, and scatterings of electrons or α-particles by an atom.
dynamics and to operate only with abstract concepts such as transition probabilities, energy levels, and the like”. Schrödinger thus emphasizes that for collision and scattering problems Heisenberg’s matrix mechanics based on an epistemology in Mach’s sense is ineffective, and his wave mechanics capable of giving intuitional images of collision and scattering processes is superior.

Seen from the three-stage theory of the process of cognizing the nature, matrix mechanics based on Heisenberg’s theory is a phenomenological theory, and wave mechanics proposed by Schrödinger who criticizes Heisenberg’s standpoint is a substantialistic theory. The three-stage theory derives from the critique of epistemologies of Mach’s type for their ineffectiveness in practice. The phenomenological stage, that is the first stage of cognizing the nature, is the stage in which the theory is based on the epistemology of Mach’s type. In the substantialistic stage, that is the second stage of cognizing the nature, the substantial structure of the object gets cognized, by means of which the knowledge of the phenomenon is arranged and the phenomenological laws are regulated. Matrix elements, that are premised in matrix mechanics as “quantities to substitute for classical quantities”, turn out to be given by the mediation of the eigenfunctions of Schrödinger’s wave equation, and the algebraic rules found for the matrix elements are derived from the properties of the set of eigenfunctions, as we have seen in the last subsections.

The theory in the substantialistic stage, however, does not assert that certain events should necessarily follow an event. In matrix mechanics, the matrix elements of the electric dipole moment of an electron can be assumed in a correspondence theoretical way to give the transition probability per unit time, on account of the introduction of matrix elements as quantities to substitute for classical quantities. In wave mechanics, on the other hand, this cannot be asserted. In order to solve this problem, it is necessary to have the equation governing the time dependence of the wave functions, and further to make clear the essence of the wave functions.

The former objective gets achieved in Schrödinger’s fourth paper, as we shall describe in Sec. 4 of the present chapter. An important step for the solution of the latter objective is made by Born’s application of the wave equation to the scattering of electrons, of which Schrödinger emphasized the necessity of having “intuitional” images in the sense of wave mechanics, that is, the substantialistic knowledge. This will be treated in Sec. 5 of the present chapter.
4. Perturbation Theory and Dispersion Formula of Light

§1. Perturbation theory for energy eigenvalues

In the third paper on "Quantization as an Eigenvalue Problem",\textsuperscript{34} received on May 10, 1926, Schrödinger develops the perturbation theory for the wave equation to treat the Stark effect of the hydrogen atom. As we have mentioned in Sec. 2, §2 of the present chapter, he presented in the second paper\textsuperscript{35} several examples of application of the wave equation. In the example of a diatomic molecule, he made a calculation of the effect of rotation on the relative oscillation in disregard of nonharmonic terms arising from the centrifugal force.

About this, Schrödinger argues in the second paper that "the method selected by Kratziel with great skill for the classical mechanical treatment is suitable also for the undulatory mechanics", but "in order to carry on the calculation to as far as is necessary for the fineness of the band structure, one must make use of the perturbation of the eigenvalues and eigenfunctions". He mentions also that "this perturbation theory is the complete counterpart to that in classical mechanics, but it is simpler for the reason that in the undulatory mechanics we are exclusively in the field of linear relations", and "the perturbation theory broadens the width of the analytical grasping extraordinarily".

Next, in the third paper, Schrödinger studies the perturbation theory for the eigenvalues of the Sturm-Liouville differential equation, citing Courant-Hilbert's famous textbook on mathematics in physics,\textsuperscript{36} and derives the first order perturbation formula for energy eigenvalues. He applies the result to the Stark effect of the hydrogen atom to get the splitting of spectral terms. He calculates also the relative intensities of spectral lines in the Stark effect on an assumption, as will be shown in the next subsection.

Schrödinger takes the linear homogeneous second order differential expression

\[ L[y] = py'' + p'y' - qy, \quad (88) \]

and treats the Sturm-Liouville eigenvalue problem

\[ L[y] + Epy = 0, \quad (89) \]

\textsuperscript{34}E. Schrödinger, Ann. d. Phys. 80 (1926), 437.
\textsuperscript{35}The same as Ref. 16) of the present chapter.
\textsuperscript{36}R. Courant und D. Hilbert, "Methoden der mathematischen Physik" (Springer, 1924).
where \( y \) is a dependent function, \( p, q \) and \( \rho \) are continuous functions of the independent variable \( x \), and \( E \) is an eigenvalue, the prime denoting differentiation with respect to \( x \). \( L \) makes a self-adjoint differential expression.

Let \( y(x) = u_j(x) \) be the eigenfunction of Eq. (89) belonging to the eigenvalue \( E_j \). The functions \( u_j(x)\sqrt{\rho(x)} \), \( j = 1, 2, 3, \ldots \), constitute an orthogonal complete set, such that

\[
\int \rho(x)u_j(x)u_k(x)dx = 0, \quad (j \neq k). \tag{90}
\]

Schrödinger considers then

\[
L[y] - \gamma y + E\rho y = 0, \tag{89'}
\]

which results from adding a “small perturbation term” to the left-hand side of Eq. (89), \( \lambda \) being “a small quantity (perturbation parameter)” and \( r \) a continuous function of \( x \). This perturbation term corresponds to a slight change of \( q \) in Eq. (88).

Denoting the eigenfunction of Eq. (89') belonging to the eigenvalue \( E^*_k \) by \( u^*_k(x) \), and putting

\[
E^*_k = E_k + \lambda \epsilon_k, \quad u^*_k(x) = u_k(x) + \lambda \nu_k(x) \tag{91}
\]

one gets, from the continuity of the eigenvalues and eigenfunctions with respect to the change in the coefficients of the differential equation, the equation

\[
L[\nu_k] + E_k\rho \nu_k = (r - \epsilon_k \rho)u_k, \tag{92}
\]

neglecting terms of the order of \( \lambda^2 \). Schrödinger says that this inhomogeneous equation, Eq. (92), has a solution only when its right-hand is orthogonal to an eigenfunction of the homogeneous equation \( L[\nu_k] + E_k\rho \nu_k = 0 \), say \( \nu_k = u_k \).

He notes that “this is just the jumping-off place for the whole perturbation theory”, citing Courandt-Hilbert’s book. For the oscillation of a string, this “mathematical law” has the “physical interpretation” that the external force does not do work on the string in the eigen oscillation under consideration, because if it does so the amplitude of the oscillation should increase unlimtedly so that no stable state would be possible.

From this condition, one has \( \int (r - \epsilon_k \rho)u_k^2dx = 0 \) and therefore

\[
\epsilon_k = \int ru_k^2dx / \int \rho u_k^2dx. \tag{93}
\]
The inhomogeneous equation (92) can be solved by expanding \( v_k \) and \{the right-hand side of Eq. (92)\}/\( \rho \) in terms of the set of the functions \( u_j \), with the result

\[
\begin{align*}
  u_k^*(x) &= u_k(x) + \lambda \sum_{j=1}^{\infty} \frac{u_j(x) \int r u_k u_j dx}{E_k - E_j}, \\
  E_k^* &= E_k + \lambda \int r u_k^2 dx,
\end{align*}
\]  

where \( \sum' \) denotes summation over \( j \), except for \( j = k \), and \( u_j \) is assumed to be normalized such that \( \int \rho u_j^2 dx = 1 \).

In the case of a partial differential equation, the problem of degeneracy should be taken into consideration, besides the multiplicity of independent variables. Let the independent variables \( x_1, x_2, \ldots \) be denoted briefly by \( x \) and the integration over \( x_1, x_2, \ldots \) be written simply as \( \int dx \). \( L[y] \) is now supposed to denote a self-adjoint second order differential expression. Equation (89) can then be regarded as the formulation of the Sturm-Liouville eigenvalue problem with multiple independent variables. If the eigenvalue \( E_k \) is degenerate \( \alpha \)-fold, there are \( \alpha \) independent eigenfunctions \( u_{k1}, u_{k2}, \ldots u_{k\alpha} \) belonging to this eigenvalue. They are assumed to be orthogonalized and normalized as

\[
\int \rho(x) u_{kj}(x) u_{k'j'}(x) dx = \begin{cases} 0 & (k, j) \neq (k', j') \\ 1 & k = k' \text{ and } j = j' \end{cases}.
\]  

Putting in place of Eq. (91)

\[
E_{kl}^* = E_{kl} + \lambda \epsilon_l, \quad u_{kl}^* = \sum_{j=1}^{\alpha} \kappa_{lj} u_{kj}(x) + \lambda \nu_l(x),
\]  

where \( l = 1, 2, \ldots \alpha \), one gets in place of Eq. (92),

\[
L[\nu_l] + E_k \rho \nu_l = \sum_{j=1}^{\alpha} \kappa_{lj} (r - \epsilon_l \rho) u_{kl}.
\]  

For the same reason as before, the right-hand side of this equation should be orthogonal to \( u_{km} \) (\( m = 1, 2, \ldots \alpha \)). From this condition one gets

\[
\sum_{j=1}^{\alpha} \kappa_{lj} \epsilon_{mj} - \kappa_{lm} \epsilon_l = 0,
\]  

where \( \epsilon_l \) is the \( l \)-th root of the equation

\[
\int \rho(x) u_{kj}(x) u_{k'j'}(x) dx = 0.
\]
where $\varepsilon_{mj}$ is the element of a "symmetric constant matrix" given by

$$
\varepsilon_{mj} = \int r u_{km} u_{kj} dx. \tag{98'}
$$

$\varepsilon_l$ is obtained with the use of the condition that the simultaneous equations (97) have solutions, that is, the vanishing of the coefficient determinant. Solving Eq. (92') by expanding $\nu_l$ and \{the right-hand side of Eq. (92')\}/$\rho$ in terms of the set of the functions $u_{k'j'}$, one gets

$$
u_l(x) = \sum_{j=1}^a \kappa_{lj} \left\{ u_{kj}(x) + \lambda \sum_{(k',j')} \frac{u_{k'j'}(x)}{E_k - E_{k'}} \int r u_{kj} u_{k'j'} dx \right\}, \tag{94'}
$$

where $\sum'$ denotes summation over $k'$ and $j'$, except for $k' = k$.

The results thus obtained by Schrödinger for the perturbation of the eigenvalues and eigenfunctions, including the case of degeneracy, are the same as the first order perturbation formulae shown in present-day textbooks on quantum mechanics. Schrödinger next applies the results to the Stark effect of the hydrogen atom. If the external electric field is of strength $E$ and in the direction of the $z$-axis, the wave equation becomes

$$
\Delta \psi + \frac{8\pi^2 m}{\hbar^2} \left( E + \frac{e^2}{r} - e F z \right) \psi = 0. \tag{99'}
$$

Schrödinger chooses rotational parabolic coordinates $(\lambda_1, \lambda_2, \phi)$, defined by

$$
x = \sqrt{\lambda_1 \lambda_2} \cos \varphi, \quad y = \sqrt{\lambda_1 \lambda_2} \sin \varphi, \quad z = (\lambda_1 - \lambda_2)/2. \tag{100}
$$

Because $z$ is proportional to the difference between $\lambda$ and $\lambda_2$, separation of variables can be applied to the wave equation inclusive of the perturbational term $-e F z \psi$, and there does not appear the problem of degeneracy such as mentioned above. He applies Eq. (95) to each of the two differential equations separated with respect to $\lambda_1$ and $\lambda_2$. In each of these two cases, the wave function is expressed in terms of the Laguerre polynomials. As a result, Schrödinger obtains for the spectral terms of the hydrogen atom in the Stark effect the formula

$$
E = -\frac{2\pi^2 m e^4}{\hbar^2 l^2} - \frac{3 \hbar^2 F l (k_2 - k_1)}{8 \pi^2 m e}, \tag{101}
$$

where $l = k_1 + k_2 + n + 1$ is the principal quantum number, $k_1$ and $k_2$ are the quantum numbers with respect to the coordinates $\lambda_1$ and $\lambda_2$, respectively, and
\( n \) is the azimuthal quantum number, \( k_1, k_2 \) and \( n \) taking the values 0, 1, 2, \ldots. This result is identical with "Epstein's well-known formula" derived by applying the quantum condition.\(^{37}\)

Schrödinger also treats the same problem using spherical coordinates \((r, \theta, \varphi)\). Though the unperturbed wave equation is separable in the spherical coordinates, as was shown in his first paper, the perturbational term is not on account of the relation \( z = r \cos \theta \), so that there arises the problem of degeneracy in the meaning described above. He calculates \( \varepsilon_{mj} \) given by Eq. (98') with the use of the wave function of the hydrogen atom in the spherical coordinates, and solves the cellular equation in the first order approximation to derive the spectral terms, arriving at a formula equivalent to Eq. (101).

As we have previously described, the quantum condition is interpreted by Schrödinger to correspond to the theory of Fresnel zones. In the case of a perturbational effect of the first order in the external electric field, however, it turns out that wave mechanics gives the same result as that derived from the quantum condition. For a second order perturbation in the external electric field, wave mechanics gives a result better than that obtained by Epstein using the quantum condition, as was shown by Wentzel by means of the WKB method in his paper\(^{38}\) received about one month later than Schrödinger's third paper under consideration.

\section*{§2. Intensity of spectral lines}

As we have mentioned in §3 of the last section, Schrödinger's theory presenting a differential equation for stationary waves cannot give the transition probability per unit time of an atom. Concerning this difficulty, Schrödinger says in his second paper\(^{39}\) that he has "the quite definite hope" that, his theory and Heisenberg's study, "these two advances do not fight each other, but rather, just because of the extraordinary difference of the starting points and the methods, will complement each other, as one helps where the other fails". He says further that "the strength of Heisenberg's programme lies in the fact,

\(^{37}\)Cf. Sec. 2, §5, Chap. 3, Vol. II. Note that Schrödinger's \( n + 1 \), instead of \( n \) itself, should be compared with Epstein's azimuthal quantum number \( n \). The same situation holds for the case of Schrödinger's derivation of the Balmer formula described in Sec. 2 of the present chapter, as well as Pauli's derivation according to matrix mechanics, which we have seen in the last chapter.

\(^{38}\)Cf. Sec. 2, §3 of the present chapter.

\(^{39}\)E. Schrödinger, Ref. 16) of the present chapter.
that it promises to give the intensity of lines, a question from which we have kept away until now”.

In his “paper on the relationship”\(^{40}\) between wave mechanics and matrix mechanics, Schrödinger gives the following arguments, trying to approach somehow this question. “The question of the coupling between the atomic dynamical event and the electromagnetic field” is particularly important, and is “perhaps the cardinal question of the whole atomic dynamics”. “The matrix representation of the atomic dynamics has led to the supposition that, actually, the electromagnetic field must be also represented differently, namely according to matrix, for the coupling to get mathematically formulated”. “The undulatory mechanics shows that about this no constraint exists anyway, because the mechanical scalar field (represented by \(\psi\)) possesses completely the aptitude to enter between the vector fields even in the unaltered Maxwell-Lorentz equations as the ‘source’ of them”. “Inversely electrodynamical potentials enter into the coefficients of the wave equation, which determines the mechanical scalar field”.

Schrödinger thinks that it is “worthwhile” to bring “a four-vector derived in an appropriate way from the mechanical scalar field \(\psi\) of the motion of an electron” into the Maxwell-Lorentz equations “as a four-current”. In the case of a many-electron problem, “the difficulty lies in the fact that \(\psi\) is a function in the configuration space, not in the actual space”. He considers, however, that if he would clarify the one-electron problem anyhow, “it should be possible to give in this way an extraordinary intuitional interpretation of the intensity and polarization of radiation”.

Thus, Schrödinger takes the case of the hydrogen atom in a state for which the “mechanical scalar field” \(\psi\) is given in the form of a series of the eigenfunctions \(u_k\) as

\[
\psi = \sum_k c_k u_k e^{2\pi i E_k t / \hbar},
\]

with real constants \(c_k\) and energy eigenvalues \(E_k\). He then makes a new assumption that the space density of electric charge is given by the real part of

\[
\psi \left( \frac{\partial \bar{\psi}}{\partial t} \right).
\]

\(^{40}\)E. Schrödinger, Ref. 27) of the present chapter.
On this assumption the space density becomes

\[
2\pi \sum_{(k,m)} c_k c_m \frac{E_k - E_m}{\hbar} u_k(x)u_m(x) \sin \left( \frac{2\pi (E_k - E_m)}{\hbar} t \right).
\] (103')

Though in the paper on the relationship Schrödinger gives no particular explanation of this assumption, it may be the reason that the time and space components of the four-current are proportional to the density and velocity of the electric charge, respectively, while in wave mechanics the momentum corresponds to \((\hbar/2\pi i)\partial/\partial x\), etc. In fact, he mentions in a footnote added in proof to the third paper, “By now I have recognized this \(\partial/\partial t\), by which I had hoped to make easier the special relativistic generalization, to be incorrect”. At the same time he states that the assumption (103) “is to be replaced by \(\psi \bar{\psi}\)”. He applies this revised assumption in his fourth paper, which we shall treat in §4 of the present section.

Now, the integration of the product of the space density of electric charge by the coordinate \(q_l\), with the use of Eqs. (102) and (103), gives for the component \(M_l\) of the electric dipole moment in the direction of \(q_l\) the expression

\[
M_l = 2\pi \sum_{(k,m)} c_k c_m q_l^{km} \frac{E_k - E_m}{\hbar} \sin \left( \frac{2\pi (E_k - E_m)}{\hbar} t \right),
\] (104)

where \(q_l^{km}\) is the matrix element given by Eq. (72) in the present chapter. This expression is “a Fourier development, in which only the differences between spectral terms occur”. Schrödinger considers that “in the coefficients there appear Heisenberg’s matrix elements in such a way that its determining influence on the intensity and polarization of the concerned part of the emitted radiation becomes completely understandable on the basis of the classical electrodynamics’.”

However, “the above sketch of the mechanism of radiation is not by any means fully satisfactory and on no account conclusive”, as Schrödinger himself points out. For example, the space density given by Eq. (103’) “gives null”, because of the orthogonality of the eigenfunctions and the factors \((E_k - E_m)\), when it is integrated over the whole space. Furthermore, “the calculation of magnetic radiation should be carried out”.

In this way, in the paper on the relationship Schrödinger made efforts to grasp, from the standpoint of wave mechanics, the supposed relation of the transition probability and the matrix element, in spite of the mentioned
difficulties. Now, in the third paper, he treats the relative intensity of the spectral lines in the Stark effect of hydrogen atom, making use of the understanding arrived at in the paper on the relationship. In order to calculate the relative intensities according to matrix mechanics, "the ratio of \( q \) must first be squared and then multiplied by the ratio of the fourth power of the frequency of the emitted radiation". But, in the Stark effect "the latter ratio plays no role, because only lines of almost equal frequencies are always compared". In this meaning the "additional factor \( (E_k - E_m)/h \)" in Eq. (104) does not come into question.

The selection rule and the polarization rule can be obtained "from the form of the eigenfunctions", without carrying out the integration to define \( q^{km}_\ell \). Because the component of the electric dipole moment in the direction of the external electric field is proportional to \( (\lambda_1 - \lambda_2)/2 \) and independent of the azimuthal angle \( \varphi \), as is seen from Eq. (100), its matrix element differs from zero only for pairs of the eigenfunctions with equal azimuthal quantum number \( n \). Since the component of the electric dipole moment perpendicular to the external field, that is in the direction of the \( x \)- or \( y \)-axis, contains the factor \( \cos \varphi \) or \( \sin \varphi \), as is seen also from Eq. (100), its matrix element has a nonvanishing value for the pairs of eigenfunctions for which \( \Delta n = \pm 1 \).

Schrödinger compares his theoretical relative intensities of \( H_\alpha \), \( H_\beta \), \( H_\gamma \), and \( H_\delta \) lines of the hydrogen atom in the Stark effect with the experimental data, as shown in the figure on the next page. In Sec. 4, §5, Chap. 3, Vol. II, we have shown a similar figure due to Kramers, who calculated in 1919 the quantities concerned on the basis of the correspondence principle. The experimental data used by Schrödinger and those adopted by Kramers are the ones observed by Stark in 1915.\(^{41}\) As Schrödinger says, "the agreement is a fairly good one for nearly all the strong components, taken all in all, certainly somewhat better than the estimation at that time from the correspondence-theoretical considerations".

So, the effectiveness of the expression in matrix mechanics to give the transition probability in terms of the matrix elements of the electric dipole moment is demonstrated by comparison with experiments by Schrödinger who relates the matrix element to the wave function. The only result on the intensity of spectral lines achieved so far in matrix mechanics is the derivation of the relative intensity of lines in the normal Zeeman effect.

From the Proposal of Wave Mechanics to Quantum Mechanics

Theoretical

Experimental

"par" and "perp" signify parallel and perpendicular components respectively.

The numbers are the values of $l (k_1 - k_2)$ of the respective transitions.
§3. Time dependence of wave function and complex wave function

Equation (51) in Sec. 2, §2 of the present chapter, which Schrödinger regarded in his second paper as the fundamental equation in wave mechanics, can be rewritten as

$$\Delta \psi + \left( \frac{8\pi^2}{\hbar^2} \right) (E - V) \psi = 0,$$

if \( \text{div} (\text{grad} \ \psi) \) is expressed as \( \Delta \psi \). Equation (51) was obtained from the equation

$$\Delta \psi + \left( \frac{1}{u^2} \right) \frac{\partial^2 \psi}{\partial t^2} = 0$$

for a wave of phase velocity \( u = E/\sqrt{2(E - V)} \), by supposing that its frequency was given by \( \nu = E/h \), in other words by putting

$$\psi \sim \text{Re}(e^{\pm 2\pi iEt/h}),$$

where \( \text{Re} \) means the real part.

Schrödinger says, in his fourth paper on “Quantization as an Eigenvalue Problem”, that Eq. (106) “is in reality not more general than” Eq. (105), because the phase velocity \( u \) is related to a particular value of the energy \( E \). “There is the need to extend the theory to nonconservative systems, because only in this way can the behavior of the system under the influence of a predetermined external force, for example, a light wave or a passing-by foreign atom, be studied”. For a time-dependent \( V \), a function which changes with the time according to Eq. (107) does not satisfy Eq. (105) or Eq. (106). Schrödinger notes that Eq. (105) or Eq. (106) is more correctly an “oscillation equation” or “amplitude equation”, but not a “wave equation” in the proper meaning.

Equation (107) is synonymous with

$$\frac{\partial^2 \psi}{\partial t^2} = - \left( \frac{4\pi^2 E^2}{\hbar^2} \right) \psi.$$  

Because one has from Eq. (105),

$$\left\{ \Delta - \left( \frac{8\pi^2}{\hbar^2} \right) V \right\}^2 \psi = \left( \frac{8\pi^2}{\hbar^2} \right)^2 E^2 \psi,$$

From the Proposal of Wave Mechanics to Quantum Mechanics

one then gets

\[ \left\{ \Delta - \left( \frac{8\pi^2}{h^2} \right) V \right\}^2 \psi + \left( \frac{16\pi^2}{h^2} \right) \frac{\partial^2 \psi}{\partial t^2} = 0. \]  

(109)

This equation holds for an “arbitrary \( E \)”. The equation “is hence obviously the standardized and general wave equation for the scalar field \( \psi \)”.

This is a fourth order differential equation with respect to the position coordinates, and is “of a type very similar to what appears in many problems of the theory of elasticity”.\(^{43}\) Schrödinger says that “about this one needs not fear excessive complication of the theory, or the necessity at all of a revision of the methods previously given in connection with” Eq. (105). If \( V \) does not contain time, one gets from Eq. (109), with the use of Eq. (107),

\[ \left\{ \Delta - \left( \frac{8\pi^2}{h^2} \right) (V - E) \right\} \left\{ \Delta - \left( \frac{8\pi^2}{h^2} \right) (V + E) \right\} \psi = 0. \]  

(109')

Schrödinger splits this equation “on a trial basis” into the following two equations,

\[ \left\{ \Delta - \left( \frac{8\pi^2}{h^2} \right) (V \pm E) \right\} \psi = 0. \]  

(110)

Because Eq. (109') is not an algebraic equation, “this splitting is not inevitable”, but “the procedure finds its later justification through a proof of completeness of the eigenfunctions as functions of coordinates”.\(^{44}\) One of Eq. (110) is identical with the equation which one gets by inverting \( E \) in Eq. (105) into \(-E\), while the other is the same as Eq. (105). This means that “the imaginary part” of \( e^{\pm 2\pi iEt/h} \) satisfies Eq. (109), as well as the real part given in Eq. (107).

Schrödinger says that the wave equation (109) “carries in itself the dispersion law”, and “can actually be regarded as the basis of the theory developed up to now of conservative systems”. The dispersion theory here means, of

\(^{43}\)Schrödinger mentions that in the case of an oscillating plate one has \( \Delta \Delta u + \partial^2 u/\partial^2 t = 0 \), citing Courant-Hilbert’s textbook mentioned before in Ref. 36) of the present chapter. In Ritz’s special oscillating plate model in his theory of spectral lines, which we have treated in Sec. 2, §1, Chap. 2, Vol. II of our present book, he used an eighth order partial differential equation with respect to the position coordinates.

\(^{44}\)For the equation \( (d^2/dx^2 + k^2)\psi = (d/dx - ik)(d/dx + ik)\psi = 0 \) with constant \( k \), for example, \( \psi = A \exp(+ikx) + B \exp(-ikx) \), given in terms of the solutions \( \psi = \exp(\pm ikx) \) of the equations \( (d/dx \pm ik)\psi = 0 \) with arbitrary constants \( A \) and \( B \), forms the complete solution. Schrödinger might hope a similar situation to hold for Eqs. (109') and (110).
course, the relation mentioned above between the phase velocity $u$ and the frequency $\nu$. Now, if one applies Eq. (109) as it is to a nonconservative system, in which $V$ changes with time, "one meets with complications which seem to come from a term involving $\partial V/\partial t".\textsuperscript{45} Hence, he "steps onto a somewhat different way", which is "by calculation much simpler" and which he thinks "in principle correct", as follows.

"One must not push up the order of the wave equation to forth in order to remove the energy parameter from it". The time dependence required for the validity of Eq. (105) can also be put as

\[
\frac{\partial \psi}{\partial t} = \pm \left( \frac{2\pi i}{h} \right) E\psi, \quad (111)
\]

instead of Eq. (108). "One arrives thus at one of the following equations".

\[
\Delta \psi - \left( \frac{8\pi^2}{h^2} \right) V\psi \mp \left( \frac{4\pi i}{h} \right) \frac{\partial \psi}{\partial t} = 0. \quad (112)
\]

"We shall demand that the complex wave function $\psi$ satisfies one of these equations". Schrödinger says that "one will be allowed to regard the real part of $\psi$ as a real wave function (if man needs it)", because the complex conjugate function $\bar{\psi}$ satisfies the other equation.

The second equation in Eq. (112), that is the one with the plus sign, is called nowadays time-dependent Schrödinger equation. At the time of its presentation, however, $\psi$ was supposed by Schrödinger to satisfy "one of the equations" given as Eq. (112) here, as we have described above. We shall return to this problem in Sec. 1, §5 of the next chapter.

In order to get a "general" wave equation from Eq. (105) by eliminating the energy $E$, Schrödinger expressed the real function $\psi$ with frequency $\nu = E/h$ as Eq. (107), with the use of complex exponential functions for the convenience of calculation as in the classical theory. This procedure brought in, however, the differential equation (109) containing complications due to $\partial V/\partial t$ for a nonconservative system. He decomposed thereupon Eq. (109) into two components, and supposed the essentially complex time dependence Eq. (111) for $\psi$ with frequency $\nu = E/h$, arriving at the time-dependent Schrödinger equation. In this way the wave equation, which is not for mathematical convenience

\textsuperscript{45} It may be that he thinks of treating effects of the time variation of $V$ in a perturbation-theoretical way by expanding $V(t)$ as $V(t) = V_0 + (\partial V/\partial t)_0 t + \cdots$. 
as in classic theory but essentially complex, was introduced for the first time together with the fundamental equation it must obey.

The frequency $\nu = E/h$ of the wave function was introduced to reproduce Bohr's frequency relation, as we have described in Sec. 2, §2 of the present chapter. The supposition newly made of the complex exponential time dependence of the wave function $\psi \sim e^{2\pi i \nu t}$ or $\psi \sim e^{-2\pi i \nu t}$, together with the essentially complex matrices and commutation relations introduced by Born-Jordan, opened the way to the unification of matrix mechanics and wave mechanics, as we shall see in Sec. 2, §1 of the next chapter. The essentially complex matrices and commutation relations are based also on Ritz's combination rule and Bohr's frequency relation, as we have seen in Sec. 2, §2 in the last chapter.

§4. **Time-dependent perturbation and dispersion formula of light**

In the fourth paper, Schrödinger makes a study of the dispersion of light, which is one of the main objects of his search of the time-dependent wave equation. He takes for the potential energy $V$ the expression

$$V = V_0(x) + A(x) \cos 2\pi \nu t, \quad (113)$$

on the assumption that a linearly polarized monochromatic light of frequency $\nu$ is incident on the atom. In this expression, $V_0(x)$ is the potential energy of the atom itself, and $A(x)$ is given, "according to the ordinary mechanics", by the negative product of the component in the direction of the polarization of the dipole moment of the atom by the amplitude of the electric field of the light.

Inserting Eq. (113) into Eq. (112), he gets

$$\Delta \psi - \left( \frac{8\pi^2}{h^2} \right) (V_0 + A \cos 2\pi \nu t) \psi + \left( \frac{4\pi i}{\hbar} \right) \frac{\partial \psi}{\partial t} = 0. \quad (114)$$

For $A = 0$, he puts

$$\psi = u(x)e^{\pm 2\pi i Et/\hbar},$$

and assumes that the eigenvalue $E_k$ and eigenfunction $u_k(x)$ of the unperturbed system have been obtained ($k = 1, 2, 3, \ldots$). For $A \neq 0$, he puts

$$\psi = u_k(x)e^{\pm 2\pi i E_k t/\hbar} + w(x, t), \quad (115)$$
and gets the inhomogeneous equation satisfied by $w$

$$
\Delta w - \left( \frac{8\pi^2}{h^2} \right) V_0 w - \left( \frac{4\pi i}{h} \right) \frac{\partial w}{\partial t}
= \left( \frac{4\pi^2}{h^2} \right) Au_k(x) \left\{ e^{2\pi it(E_k + h\nu)/h} + e^{2\pi it(E_k - h\nu)/h} \right\}.
$$

(116)

In consideration of the form of the inhomogeneous terms in the right-hand side of this equation, he puts further

$$
w(x, t) = w_+(x) e^{2\pi it(E_k + h\nu)/h} + w_-(x) e^{2\pi it(E_k - h\nu)/h},
$$

(117)

to derive

$$
\Delta w_\pm + \left( \frac{8\pi^2}{h^2} \right) (E_k \pm h\nu - V_0) w_\pm = \left( \frac{4\pi^2}{h^2} \right) Au_k.
$$

(118)

He applies the perturbation theory given in his third paper (cf. §1 of the present section) to this equation to get $w_\pm$ by expanding them in terms of the eigenfunctions $u_k$, getting the

$$
\begin{cases}
  w_\pm(x) = \frac{1}{2} \sum_{n=1}^{\infty} \frac{a_{kn} u_n(x)}{E_k - E_n \pm h\nu}, \\
  a_{kn}' = \int A(x) u_k(x) u_n(x) \rho(x) dx.
\end{cases}
$$

(119)

(120)

Here, it is assumed that $h\nu \neq |E_k - E_n|$. For $h\nu = |E_k - E_n|$ there arises a “resonance”. In the case of off-resonance, $\psi$ is consequently found to be

$$
\psi = u_k(x) e^{2\pi i E_k t/h}
+ \frac{1}{2} \sum_{n=1}^{\infty} a_{kn}' u_n(x) \left\{ \frac{e^{2\pi i(E_k + h\nu)t/h}}{E_k - E_n + h\nu} + \frac{e^{2\pi i(E_k - h\nu)t/h}}{E_k - E_n - h\nu} \right\}.
$$

(121)

Now, “in order to know the meaning of our result for the dispersion theory”, Schrödinger assumes that $\psi \bar{\psi}$ calculated from this $\psi$ gives the density of electric charge. As he says in a footnote in the paper on the relationship, “instead of $\psi \bar{\psi}$ the real part of $\psi \bar{\psi}$ was proposed”, but “it was a wrong choice, as it was already improved in the third paper”. “According to the heuristic hypothesis of the electrodynamic meaning of the scalar field $\psi$, $\psi \bar{\psi}$ represents the density of electric charge when the one-electron problem is treated”. Schrödinger
From the Proposal of Wave Mechanics to Quantum Mechanics

considers, as a "logical generalization of this hypothesis", that the density of electric charge of a mass point corresponds in wave mechanics to "the classical charge of the mass point in question, multiplied by the integral of \( \psi \bar{\psi} \) over all those coordinates of the system that fix in classical mechanics the positions of the other mass points". "The total charge density at a spatial point is then represented by the sum of the integrals mentioned taken over all the mass points".

According to this hypothesis the \( y \)-component, for example, of "the wave mechanical dipole moment" is given by

\[
\int M_y \psi \bar{\psi} \rho \, dx = b_{kk} + 2F \cos 2\pi \nu t \sum_{n=1}^{\infty} \frac{(E_n - E_k)a_{kn}b_{kn}}{(E_k - E_n)^2 - (h\nu)^2},
\]

(122)

where

\[
a_{kn} = \int M_z u_k(x) \rho(x) \, dx, \quad M_z = \sum e_i z_i,
\]

(123)

\[
b_{kn} = \int M_y u_k(x) \rho(x) \, dx, \quad M_y = \sum e_i y_i.
\]

(124)

In this calculation, the incident light is assumed to be polarized in the \( z \)-direction with the amplitude \( F \) of the electric field, and only terms up to the first order in the perturbed part of \( \psi \) or \( \bar{\psi} \) are retained.

Schrödinger supplements his analyses of an unperturbed system with eigenvalues ranging continuously in the interval \( a \leq E \leq b \), showing that in this case the right-hand side of Eq. (122) should have the the additional term

\[
2F \cos 2\pi \nu t \int_a^b \frac{(E - E_k)\alpha_k(E)\beta_k(E)}{(E_k - E)^2 - (h\nu)^2} \, dE,
\]

(122')

where

\[
\alpha_k(E) = \lim_{\Delta \to 0} \frac{1}{\Delta} \int \rho(\xi)M_z(\xi)u_k(\xi) \int_{E}^{E+\Delta} u(\xi, E')dE'd\xi,
\]

(123')

\[
\beta_k(E) = \lim_{\Delta \to 0} \frac{1}{\Delta} \int \rho(\xi)M_y(\xi)u_k(\xi) \int_{E}^{E+\Delta} u(\xi, E')dE'd\xi,
\]

(124')
and \( u(x, E) \) denotes the eigenfunction belonging to the continuous eigenvalue \( E \), with the orthonormalization

\[
\int dx \rho(x) \int_{E'}^{E'+\Delta} u(x, E)u(x, E'')dE'' = \begin{cases} 1 & (E' \leq E \leq E' + \Delta) \\ 0 & \text{(otherwise)} \end{cases}.
\] (125)

The dipole moment given above in Eq. (122) "is to be attributed to the secondary radiation". Schrödinger says that "its term variable in time looks quite reasonable and might meet the requirements which one is usually to put on a 'dispersion formula'". "One over all observes the appearance also of the so-called 'negative' terms which correspond to the possibility of a transition to a deeper level \( (E_n < E_k) \) and which Kramers has first directed the attention to on the basis of correspondence-theoretical considerations". "In general our formula is — in spite of the very different denotation and way of thinking — well describable as formally identical with Kramer's formula of secondary radiation". "The important connection of the coefficients of secondary radiation with the coefficients of spontaneous radiation \( a_{kn} \) and \( b_{kn} \) is put in evidence, and will describe precisely the secondary radiation also with regard to its state of polarization".\(^{46}\)

As we have seen in §2 of the present section, Schrödinger previously gave, by means of the wave function, a substantialistic meaning to the matrix element, borrowed the relation between the matrix element and the transition probability, which was correspondence-theoretically assumed in matrix mechanics, by interpreting it on the assumption that \( \psi(\partial \tilde{\psi}/\partial t) \) represents the electric charge, density in space and reproduced fairly well the relative intensities of the spectral lines in the Stark effect of hydrogen atom.

In this place of the fourth paper, Schrödinger calculated the electric dipole moment from the wave function \( \psi \) perturbed by the oscillating electric field of the incident light, on the new assumption that \( \psi \tilde{\psi} \) would be proportional to the density of electric charge, getting a result corresponding just to the dispersion formula which Kramers obtained on the basis of the correspondence principle. In Schrödinger's new result, the matrix elements \( a_{kn} \) and \( b_{kn} \), which were previously borrowed from matrix mechanics as the determinatives of the spontaneous transition probability, appeared from only the time-dependent Schrödinger equation and the assumed meaning of \( \psi \tilde{\psi} \). This fact is to be

\(^{46}\) Cf. Eq. (49) due to Kramers-Heisenberg in Sec. 4, §2, Chap. 5, Vol. II. Schrödinger's \( a_{kn} \) and \( b_{kn} \) correspond to the \( z \)- and \( y \)-components, respectively, of Kramers-Heisenberg's \( A_a \) (or \( \tilde{A}_a \)) when \( E_n > E_k \), and \( A_e \) (or \( \tilde{A}_e \)) when \( E_n < E_k \).
noticed, because it shows that wave mechanics was about to get rid of its substantialistic limitation by means of the time dependent Schrödinger equation and the above interpretation as essential theoretical elements. We shall discuss further this subject in the following section.

5. The Physical Meaning of Wave Function

§1. \( \psi \bar{\psi} \) as "the weight function" and the continuity equation

As we have seen in the last subsection, in the fourth paper under consideration Schrödinger supposes, by generalizing "the heuristic hypothesis of the electrodynamical meaning of the scalar field \( \psi \) used for the one-electron problem", that "the charge density at a spatial point" is obtained in the following way. Namely, he multiplies the electric charge of a mass point by the integral of \( \psi \bar{\psi} \) carried over the coordinates of the remaining mass points, and equates the coordinates of the chosen mass point with those of the spatial point at which the charge density is to be known. The sum of the quantities so calculated taken over all the mass points must then be the charge density of the system at the spatial point in question.

Now Schrödinger says in the last section "On the physical meaning of the scalar field" of the fourth paper,\(^{47}\) that "this rule is synonymous with the view that \( \psi \bar{\psi} \) is a kind of weight function in the configuration space of the system", and that "this view lets the original meaning of \( \psi \) become better apparent". He considers that "the wave-mechanical configuration of the system is a superposition of many, strictly speaking all, kinematically possible configurations in the mechanics of mass points", and "every point-mechanical configuration contributes with a certain weight to the true wave mechanical configuration, which is just given by \( \psi \bar{\psi} \).

"If one likes paradoxes, one can say that the system turns up, so to speak, in all the mechanically possible positions at the same time, but not in all with equal strength". "In macroscopic motions the weight function contracts practically over a small region of practically undistinguishable positions, whose center of mass covers macroscopically perceptible distances in the configuration space". "Anyway in problems of microscopic motion, and for certain problems even primarily, its variable distribution across the region is of interest".

\(^{47}\)E. Schrödinger, Ref. 42) in Sec. 4, §3 of the present chapter.
After having given the new interpretation of "weight function" to \( \psi \bar{\psi} \), Schrödinger mentions that "this reinterpretation at the first sight may give a shock (mag choquieren)". This is "because until now we have often spoken of the \( \psi \)-oscillations' in so intuitionally concrete form as of something quite real". He considers, however, that "something tangibly real is based on them also according to the present view, namely the extremely real, electrodynamically effective fluctuation of the electric spatial density". He says also that "the \( \psi \)-function should neither more nor less afford, as what it allows, the whole of these fluctuations to be ruled and to be looked out over mathematically by a unique differential equation".

Of the "weight function" in the above sense, the integral carried over the whole configuration space should subject to "the persistence of the normalization" on being put equal to a definite value, desirably to unity. The electric charge of the system would otherwise not be conserved. "The persistence of the normalization" is to be guaranteed by Eq. (112), which gives the time dependence of \( \psi \). "If it were not the case, this would be quite catastrophic for our whole view". Schrödinger shows that it is really the case, as follows.

He expresses Eq. (112) as

\[
\begin{align*}
\frac{\partial \psi}{\partial t} &= \left( \frac{\hbar}{4\pi i} \right) \left( \Delta - \frac{8\pi^2 V}{\hbar^2} \right) \psi, \\
\frac{\partial \bar{\psi}}{\partial t} &= -\left( \frac{\hbar}{4\pi i} \right) \left( \Delta - \frac{8\pi^2 V}{\hbar^2} \right) \bar{\psi},
\end{align*}
\]

(126)

and multiplies the first by \( \rho \bar{\psi} \) and the second by \( \rho \psi \) to be added together, getting\(^{48}\)

\[
\frac{\partial (\rho \psi \bar{\psi})}{\partial t} = \left( \frac{\hbar}{4\pi i} \right) \rho (\bar{\psi} \Delta \psi - \psi \Delta \bar{\psi}).
\]

(127)

Because \( \rho \Delta \) is expressed in terms of the coordinates \( q_l \) as\(^{49}\)

\[
\rho \Delta = \sum_l \frac{\partial}{\partial q_l} \left\{ \rho T_{pl} \left( q_k, \frac{\partial}{\partial q_k} \right) \right\},
\]

(128)

\(^{48}\) \( \rho \) is the "density function" which is characteristic to a system of curvilinear coordinates, and is the same as \( \rho (x) \) in Eq. (65), or \( \Delta_p^{-1/2} \) in Eq. (82) of the present chapter.

\(^{49}\) \( T_{pl} \) is given by Eq. (85) in Sec. 3, §2 of the present chapter.
Eq. (127) can be put into the form
\[
\frac{\partial}{\partial t} (\rho \psi \bar{\psi}) = \frac{\hbar}{4\pi i} \sum_k \frac{\partial}{\partial q_k} \left\{ \rho \bar{\psi} T_{pl} (q_k, \frac{\partial \psi}{\partial q_k}) - \rho \psi T_{pl} (q_k, \frac{\partial \bar{\psi}}{\partial q_k}) \right\}.
\] (129)

The right-hand side of this equation is the divergence of a multi-dimensional real vector. This shows that the integral of $\rho \psi \bar{\psi}$ taken over the whole configuration space is independent of time. Schrödinger says that the multi-dimensional vector in the right-hand side of Eq. (129) “is obviously to be interpreted as the stream density in the configuration space”, and therefore Eq. (129) is “the continuity equation of the weight function”.

From Eq. (129) one can obtain “the continuity equation of the electric charge” and “that for the charge density coming from each individual mass point”. Taking into consideration the $\alpha$th mass point, one has from Eq. (129)
\[
\frac{\partial}{\partial t} \left( e_{\alpha} \int \psi \bar{\psi} dx' \right) = \frac{\hbar e_{\alpha}}{4\pi i m_{\alpha}} \text{div}_{\alpha} \left\{ \int (\bar{\psi} \text{grad}_{\alpha} \psi - \psi \text{grad}_{\alpha} \bar{\psi}) dx' \right\},
\] (130)

where $e_{\alpha}$ and $m_{\alpha}$ denote its electric charge and mass respectively, $\text{div}_{\alpha}$ and $\text{grad}_{\alpha}$ are operators with regard to its Cartesian coordinates $x_{\alpha}$, $y_{\alpha}$ and $z_{\alpha}$, and $x'$ represents in short the Cartesian coordinates of all the mass points except itself.\(^{50}\) This equation is the continuity equation for the charge density due to the $\alpha$th mass point. The sum of this equation taken over all the mass points gives the continuity equation of the charge density of the system.

The continuity equation of “the weight function” given here by Schrödinger is of the same form as the continuity equation of the probability density in present-day quantum mechanics. “The weight function” $\psi \bar{\psi}$, with its integral of normalization taken over the whole configuration space equal to unity, plays in the continuity equation the same role as the probability density in quantum mechanics. Schrödinger’s expression $\int M_y \psi \bar{\psi} \rho dx'$, given in Eq. (122) for the electric dipole moment, is, when the term “weight function” is replaced by the term probability, identical to the expectation value of $M_y$ in present-day quantum mechanics. His words quoted above “If one likes paradoxes, one can say that the system turns up, so to speak, in all the mechanically possible positions at the same time, but not in all with equal strength”, resemble closely the explanation of the probability density seen in present-day textbooks on quantum mechanics.

\(^{50}\)In the Cartesian coordinates the density function $\rho$ is equal to unity.
As is indicated by Schrödinger's words made in giving $\psi \bar{\psi}$ the new meaning of "weight function", that "this reinterpretation at the first instant may give a shock", he presented in this reinterpretation a new concept by going beyond his original idea of $\psi$. In his treatments of the wave equation by analogy with the oscillation equation of a string, and of the wave mechanics by analogy with the wave optics, Schrödinger spoke of $\psi$ "in so intuitionally concrete form as of something quite real", namely he discussed it on the basis of a substantialistic image, as we have described in detail in the present chapter. At the same time, Schrödinger stressed the superiority of the substantialistic wave mechanics over the phenomenological matrix mechanics.

However, though wave mechanics gave a substantialistic meaning to the matrix elements, it could not derive the relation of the matrix elements of electric dipole moment to the transition probability, which the matrix mechanics could premise as a phenomenological theory. Thus, Schrödinger supposed, after trial and error, the proportionality of $\psi \bar{\psi}$ to the charge density by setting up the time-dependent wave equation, arriving at a result corresponding to the dispersion formula of Kramers.

As a result of reinterpretating $\psi \bar{\psi}$ as weight function, instead of assuming its proportionality to the charge density, $\psi$ came to be grasped as the function that "should afford the whole of the tangibly real fluctuations of the electric spatial density, to be ruled and looked out over mathematically by a unique differential equation", having got rid of its character of a classical wave considered in the process of setting up the wave equation. By "fluctuation" Schrödinger meant that, while in macroscopic motions the weight function would "contract practically over a small region", in microscopic motions it would have a "variable distribution across the region", that is, the characteristic nature of quantum-theoretical motions.

Thus, Schrödinger made progress and nearly arrived at quantum mechanics. For this it would be necessary to replace the concept of "weight function" with that of probability density which is, as a matter of form, equivalent to the former. The interpretation that $\psi \bar{\psi}$ should represent the probability density was given by Born, as we shall see in the following subsections.

By the way, Schrödinger showed in his paper "The exact Transition from the Micro- to Macromechanics"\textsuperscript{51} that if for a harmonic oscillator a suitable

\textsuperscript{51}E. Schrödinger, Naturwiss. 28 (1926), 664. The date of its reception is not shown. The date of the issue is July 9, 1926. In this paper Schrödinger cited his first and second papers and the paper on the relationship already published, and mentioned his third and fourth
superposition of the eigenfunctions with high quantum numbers and small relative difference in quantum number is made to form a “wave group”, it is possible to contract its $\psi$ in the neighborhood of the position of the mass point, which oscillates according to classical mechanics, and expected that a similar situation might be realized also in the case of hydrogen atom.

§2. Born’s particle-theoretical interpretation of wave function

Born presented the outline of his study on collision problems in “For the Quantum Mechanics of Collision Processes [Provisional Communication]”, received on Jun. 25, 1926, and developed the theory fully in his paper “Quantum Mechanics of Collision Processes”, received on Jul. 21, 1926. Born’s provisional communication was received four days later than Schrödinger’s fourth paper. In Born’s full paper citation was given of Schrödinger’s first and second papers, the paper on the relationship, and the short paper given in Ref. 51) above.

About atomic or molecular collisions and scatterings of electrons or $\alpha$-particles by atoms, Schrödinger argued in the paper on the relationship, as we have described in Sec. 3, §3 of the present chapter, that “it seems extraordinarily difficult to tackle problems of this kind, so long as one feels obliged, on epistemological grounds, to suppress the intuition in the atomic dynamics and to operate only with abstract concepts such as transition probabilities, energy levels, and the like”.

Born starts his full paper mentioned above, by saying that “collision processes not only have provided the most convincing experimental evidence for the basic assumptions of the quantum theory, but also seem to be suitable to give elucidation to the physical meaning of the formal laws of the so-called ‘quantum mechanics’”, pointing out the experimental as well as theoretical significance of collision problems. One understands easily Born’s remark when one remembers the Rutherford experiment, experiments on the velocity loss of $\alpha$-particles, the Frank-Hertz experiment, the detection of electrons with counters or cloud chambers in Compton effect experiments, etc., which we have treated in Vols. I and II.

\[\text{paper as being “in press”}.\]

\[52)\text{M. Born, Zeit. f. Phys. 37 (1926), 863.}\]

\[53)\text{M. Born, Zeit. f. Phys. 38 (1926), 803.}\]
He says, in the provisional communication, that in the quantum-mechanical calculation of the interaction of two systems, "all the states of both the systems couple each other in complicated ways". This holds for any "aperiodic process such as collision". In a collision process a particle, say an electron, comes from the infinity and goes away again to the infinity. "When the electron leaves far enough and the coupling is small, a determined straight uniform motion of the electron must be definable". "The important thing is to grasp mathematically this asymptotic behavior of the coupled particle". Born says "With the matrix form of quantum mechanics I have not succeeded in doing this, but I have with Schrödinger's formulation".

Born then argues that "a straight moving electron corresponds to a plane wave", and when the electron and the atom come into interaction "there occurs a complicated oscillation". "One can however fix this by means of its asymptotic behavior at the infinity". "One has really nothing but a diffraction problem, in which an incoming plane wave is bent or scattered; in place of the boundary condition which one uses in optics for the description of a shield, one has here the potential energy of the interaction between the atom and the electron". This statement of Born may be due to his way of the treating the problem, in which only the relation between the amplitude of the incident plane wave and that of an asymptotic plane wave is taken into consideration, as will be described below.

Now, in the full paper, Born says about matrix mechanics and wave mechanics as follows. "The matrix form of quantum mechanics starts from the idea that an exact description of processes in space and time is in general impossible, and hence contents itself with the establishment of relations between observable quantities, which can be interpreted as characteristics of motions only in the classical limiting case". "On the other hand, Schrödinger appears to attribute to the waves which he regards as the carriers of atomic processes following de Broglie's idea, a reality of such kind as light waves possess; he attempts to build wave groups which have small dimensions in all the directions and should evidently directly represent moving particles".

However, "neither of the two views seems satisfactory" to him. He thus takes up the problem of attempting "to give a direct interpretation and to test its usability on collision processes". For this purpose he "takes up Einstein's remark on the relationship of wave field and light quanta". According to Born, "Einstein said roughly that waves exist only to show the corpuscular light quanta the way, and in this sense Einstein spoke of a 'ghost field'". Born says
that “this field determines the probability that a light quantum, the carrier of energy and momentum, follows a definite way, but to the field itself neither energy nor momentum belongs”.

Born considers, however, that “one will probably better postpone the putting of these thoughts in direct connection with quantum mechanics, until the arrangement of the electromagnetic field in this formalism is carried out”. On the other hand, he thinks that “on a complete analogy between light quantum and electron, one will mind to formulate the laws of motion of electrons in a similar way”, and “here it is clear that de Broglie-Schrödinger’s waves are regarded as the ‘ghost wave’, or better ‘guiding wave’”.

Born thus pursues “as trial” the idea that “the guiding field represented by a scalar function \( \psi \) of the coordinates of all the particles concerned and the time, spreads according to Schrödinger’s differential equation”. It is supposed that “momentum and energy can be transferred, when particles (electrons) actually fly around”, and “the orbits of these particles are only determined as far as the energy-momentum law restricts them”. Born says that “one could somewhat paradoxically summary this as that the motion of a particle obeys probability laws, but the probability itself spreads in accordance with the causality law”.

Einstein’s remark quoted by Born reminds us of the original purpose of de Broglie’s wave theory. As we have seen in Sec. 2, §3, Chap. 5, Vol. II, de Broglie arrived, in the course of attempting to found the theory of radiation on the basis of the light quantum theory, at the idea of the phase wave “associated with” a light quantum, and supposed that a light quantum “follows the ray of the phase wave, that is to say, the normal to the surface of equal phase”. He conceived that “the phase wave is to guide the transfer of energy, and is what can permit the synthesis of the undulation and the quantum”.

Being thorough about the light quantum-theoretical viewpoint, de Broglie came to suppose that a light quantum would have a quite small but finite rest mass. This supposition brought him the concept of the phase wave, as we have described before in Vol. II. He placed great importance to the fact that, though the phase wave was assumed to have a phase velocity greater than the light velocity, its group velocity, namely the velocity of a wave group, became equal to that of a light quantum which is slightly smaller than the light velocity. Because this relation between phase velocity and group velocity would hold for a particle with an arbitrary rest mass, de Broglie introduced the phase wave also for electron.
Since light is understood in classical theory by means of the electromagnetic wave theory, in order to achieve “the synthesis of the undulation and the quantum” by introducing another kind of wave, named “the phase wave associated with a light quantum”, a double step, so to speak, would be needed, namely both the relationship between a light quantum and the phase wave and that between the phase wave and the electromagnetic wave would come into question. On the other hand, since the electron is regarded classically as a particle, de Broglie’s theory came to the fruition of Schrödinger's monistic wave theory, on the basis of the equality of the velocity of an electron to the group velocity of the phase wave.

In Schrödinger’s wave mechanics, however, de Broglie’s very idea of the phase wave to guide a quantum disappeared, and the particle nature was grasped by means of the agreement of “geometric mechanics” as an approximation to wave mechanics with the Hamilton mechanics. Schrödinger, on the other hand, made a number of attempts to derive in a substantialistic way, from the wave function and wave equation, the relation between the matrix element and the transition probability, which matrix mechanics was able, on account of its phenomenological nature, to introduce, and finally reached the reinterpretation that $\psi \overline{\psi}$ gives “the weight function” of distribution of the representative point of the system in the configuration space. This means that he was going away from his previous monistic wave theory, as is clear from his own words that “this reinterpretation may give a shock”. We have seen that “the weight function” is mathematically identical with the probability density in quantum mechanics. It was the understanding of the particle nature, not being grasped by means of the approximation to the wave nature, that was lacking in Schrödinger’s reinterpretation of $\psi \overline{\psi}$.

In Born’s work under consideration, this lack was supplemented with the use of the analogy to the relationship, suggested by “Einstein’s remark”, between the wave- and particle-natures of light. With regard to “Einstein’s remark” Born wrote, as we have described above, that “the ghost field determines the probability that a light quantum, the carrier of energy and momentum, follows a definite way”. In this relation, one may remember that in Einstein’s paper in 1916 on the introduction of the transition probability, he “entrusted time and direction of the elementary process to ‘chance’”, as has been treated in Sec. 3, §3 and §4, Chap. 3, Vol. II.
§3. Coefficient in eigenfunction expansion and the frequency of occurrence of a state

Now, in the full paper cited above, Born makes a "mathematical development" for the sake of "preciseness of concepts". For the aid for "the mathematical practice" in this paper, Born expresses his thanks to N. Wiener in a footnote. As we have seen in Sec. 4, §2 of the last chapter, Born together with Wiener treated the problem that the matrix mechanics would hardly be applied to translational motion of a particle, and attempted to formulate matrix mechanics using the method of operators.

Born considers first a periodic system to relate $|\psi|^2$ to the probability. Let $\psi_n(q)$ be an eigenfunction of Schrödinger's differential equation, normalized according to

$$ \int \psi_n(q) \psi^*_n(q) dq = \delta_{nm}, \quad (131) $$

where the asterisk means complex conjugate and no degeneracy is assumed. Any function $\psi(q)$ can be expanded in terms of the set of the eigenfunctions as

$$ \psi(q) = \sum_n c_n \psi_n(q). \quad (132) $$

Born says that the aforementioned idea "suggests the thought that this superposition of functions is connected with the probability that in a crowd of identical and independent atoms the states appear with a definite frequency of occurrence".

"The relation of completeness,"

$$ \int |\psi(q)|^2 dq = \sum_n |c_n|^2, \quad (133) $$

leads us to regard this integral as the number of the atoms". Born considers that "because it has for the occurrence of an individual normalized eigen-oscillation the value 1 (or the a priori weights of the states are 1), $|c_n|^2$ represents the frequency of occurrence of the state $n$, and the total number is additively composed of these parts".

---

54) M. Born, Ref. 53) in §2. of the present section.
"In order to justify this interpretation", Born assumes the Schrödinger equation for a mass point

$$\Delta \psi + \left( \frac{8\pi^2 \mu}{\hbar^2} \right) (W - U) \psi = 0,$$  \hspace{1cm} (134)

where $\mu$, $W$ and $U$ are its mass, energy and potential energy respectively. For the eigenfunction $\psi_n$ with the eigenvalue $W_n$ of this equation, one has

$$\oint \oint \oint \psi_n^* \left\{ \Delta \psi_n + \left( \frac{8\pi^2 \mu}{\hbar^2} \right) (W_n - U) \psi_n \right\} d\tau = 0, \quad (d\tau = dx dy dz) \hspace{1cm} (135)$$

from which one gets, making use of Green's theorem,

$$W_n = \oint \oint \oint \left\{ \left( \frac{\hbar^2}{8\pi^2 \mu} \right) |\text{grad} \psi_n|^2 + U |\psi_n|^2 \right\} d\tau. \hspace{1cm} (136)$$

"Every energy level can therefore be grasped as the spatial integral of the energy density of the eigen-oscillation".

"Forming for some function $\psi$ the corresponding integral"

$$W = \oint \oint \oint \left\{ \left( \frac{\hbar^2}{8\pi^2 \mu} \right) |\text{grad} \psi|^2 + U |\psi|^2 \right\} d\tau, \hspace{1cm} (137)$$

one has, with the use of the expansion (132),

$$W = \sum_n |c_n|^2 W_n. \hspace{1cm} (138)$$

In accordance with the meaning of $|c_n|^2$ mentioned above, "the right-hand side is the mean value of the energy of the system of atoms".

Born considers, however, that "anything else than non-essentials can be said for our approach, as long as one remains within periodic processes", and treats aperiodic processes. In the case of a "straight uniform motion" along the $x$-axis, the differential equation concerned reads

$$\frac{d^2 \psi}{dx^2} + k^2 \psi = 0, \quad k^2 = \left( \frac{8\pi^2 \mu}{\hbar^2} \right) W. \hspace{1cm} (139)$$

For a positive value of $W$, this has the eigenfunction $\psi(k, x) \propto e^{\pm ikx}$. "In order to be able to define the weight and frequency of occurrence, one must above
all normalize the eigenfunctions". Putting

$$
\lim_{a \to \infty} \frac{1}{2} \int_{-a}^{+a} |\psi(k, x)|^2 dx = 1 \quad (140)
$$

by analogy with Eq. (131), one gets as the normalized eigenfunction

$$
\psi(k, x) = e^{\pm ikx}, \quad (141)
$$

If "one considers free motion as the limiting case of a periodic motion, namely the eigen-oscillation of a finite part of the x-axis", the number of the eigen-oscillations per unit length and per interval \((k, k + \Delta k)\) is given by \(\Delta(1/\lambda) = \Delta k/2\pi\) in terms of the wavelength. Therefore, putting

$$
\psi(x) = \int_{-\infty}^{+\infty} c(k)\psi(k, x) d(k/2\pi), \quad c(-k) = c^*(k), \quad (142)
$$

"one expects that \(|c(k)|^2\) is a measure of the frequency of occurrence for the interval \(dk/2\pi\)". By analogy with the completeness relation Eq. (133), one has

$$
\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = \int_{-\infty}^{+\infty} dx \left| \int_{-\infty}^{+\infty} c(k)e^{ikx} dk/2\pi \right|^2. \quad (143)
$$

In the case that \(k\) is restricted to the small interval \(k_1 \leq k \leq k_2\), this becomes

$$
\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = |\bar{c}|^2 \frac{(k_2 - k_1)}{2\pi}, \quad (144)
$$

where \(\bar{c}\) is a suitable mean value of \(c(k)\) in the interval.

"One can grasp this as a 'matrix'." Putting the matrix element of the momentum\(^{55}\)

$$
p(k, k') = \frac{\hbar}{2\pi i} \lim_{a \to \infty} \frac{1}{2a} \int_{-a}^{+a} \psi^*(k, x) \frac{\partial \psi(k', x)}{\partial x} dx, \quad (145)
$$

one gets

$$
p(k, k') = \begin{cases} 
(h/2\pi)k & \text{(for } k = k'\text{).} \\
0 & \text{(for } k \neq k'\text{).}
\end{cases}
$$

\(^{55}\)The relation of the matrix element to the wave functions was given in Schrödinger’s paper on the relationship, which Born cited in his paper under consideration.
Because the momentum $p$ of a translational motion is given by $p = h/\lambda = (h/2\pi)k$ according to de Broglie, one has, with $\Delta k = k_2 - k_1 = (2\pi/h)\Delta p$,

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = \frac{|c|^2 \Delta p}{h}. \quad (146)$$

This shows that "a cell of extension in length $\Delta x = 1$ and extension in momentum $\Delta p = h$ has the weight 1, and $|c(k)|^2$ is the frequency of occurrence of a motion with the momentum $p = (h/2\pi)k$".

In a collision process, "the particle is before and after the collision in a practically free state". "One comes therefore, in accordance with the experimental setting up of the problem, to the view that for the asymptotic motion before the collision the distribution function $|c(k)|^2$ should be known". "Can one calculate from it the distribution function after the collision?" Born makes this calculation by making use of the wave equation as follows.

### §4. Wave equation and scattering probability of a particle

For a motion along the $x$-axis under the influence of a potential energy $U(x)$, Schrödinger's differential equation can be put, instead of Eq. (139), in the form

$$\frac{d^2 \psi}{dx^2} + k^2 \psi = V(x), \quad V(x) = \frac{8\pi^2 \mu}{h^2} U(x). \quad (147)$$

Born applies the method of iterative approximation to solve this equation, starting from the function $u_0(x) = e^{ikx}$ on the assumption that $|V(x)| < K/x^2$, $K$ being a positive constant. He shows that in terms of the functions $u_n(x)$ determined by

$$\frac{d^2 u_n}{dx^2} + k^2 u_n = V u_{n-1} , \quad (148)$$

$\psi(x)$ is given by

$$\psi(x) = \sum_{n=0}^{\infty} u_n(x).$$

Because $|u_n(x)| \leq (K/k|x|)^n/n!$, for $x \to \pm \infty$ there exists a solution with the asymptotic form $e^{-ikx}$ besides the one with the asymptotic form $e^{ikx}$. "We are interested in the case that from $x = +\infty$ only a wave is incoming".
Therefore, with the phase velocity $u$, $\psi(x)$ at $x = +\infty$ takes the form

$$
\psi^+(x) = c_e e^{ik(x + ut)} + c_a e^{-ik(x - ut + \phi_a)},
$$

(149a)

and at $x = -\infty$ the form

$$
\psi^-(x) = C_a e^{ik(x + ut + \phi_a)}.
$$

(149b)

The first term in Eq. (149a) represents the “incoming wave”, and the second term in Eq. (149a), as well as Eq. (149b), represents the “outgoing wave”. Born says that the relations of $c_a$ and $C_a$ to $c_e$ can be obtained “very simply with the aid of the energy law”.

That is, Born assumes that $\psi$ obeys the time-dependent wave equation

$$
\frac{\partial^2 \psi}{\partial x^2} - \left( \frac{1}{u^2} \right) \frac{\partial^2 \psi}{\partial t^2} = 0,
$$

from which one can derive Schrödinger’s differential equation by putting $u^2 = W^2/2\mu(W - U)$.

If $u$ is independent of time, one gets from Eq. (150)

$$
\frac{\partial}{\partial x} \left( \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial t} \right) - \frac{\partial}{\partial t} \left\{ \frac{1}{2} \left( \frac{\partial \psi}{\partial x} \right)^2 + \frac{1}{2u^2} \left( \frac{\partial \psi}{\partial t} \right)^2 \right\} = 0.
$$

(151)

Integrating this over space, one obtains

$$
\left[ \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial t} \right]_{-\infty}^{+\infty} - \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} \frac{1}{2} \left\{ \left( \frac{\partial \psi}{\partial x} \right)^2 + \frac{1}{u^2} \left( \frac{\partial \psi}{\partial t} \right)^2 \right\} dx = 0.
$$

(152)

Because the integral in the second term “means the total energy existing in the space”, the time average of its time derivative becomes zero. One has therefore, from the time average of the first term,

$$
\frac{\partial \psi^+}{\partial x} \frac{\partial \psi^+}{\partial t} = \frac{\partial \psi^-}{\partial x} \frac{\partial \psi^-}{\partial t} + \frac{\partial \psi^+}{\partial x} \frac{\partial \psi^-}{\partial t} - \frac{\partial \psi^-}{\partial x} \frac{\partial \psi^+}{\partial t}.
$$

(153)

This equation together with Eq. (149) gives the result

$$
c_e^2 = c_a^2 + C_a^2.
$$

(154)

---

56) This procedure is identical with that used in Schrödinger’s second paper, as we have seen in Sec. 2, §2 of the present chapter.

57) Born’s $(\partial \psi/\partial x)(\partial \psi/\partial t)$ corresponds to the quantity, in the case of a string which was often used as analogy in Schrödinger’s papers, proportional to the power at a point of the string.
This means that "for every elementary wave of given \( k \), the incoming intensity is split into the intensities of two waves scattered to the right and the left". "In the words of the corpuscular theory, when a particle with a given energy hits on the atom, it is either reflected or goes forth, the sum of the probabilities of these two events being 1". Born concludes thus that "the law of energy conservation leads to the conservation of the number of particles".

For three-dimensional motion, one can show in a similar way from the wave equation

\[
\Delta \psi - \left( \frac{1}{u^2} \right) \left( \frac{\partial^2 \psi}{\partial t^2} \right) = 0
\]  

that the following relation holds for the surface integral at the infinity,

\[
\int_{\infty} \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial t} dS = 0.
\]  

If the general solution to the differential equation for an inertial motion

\[
\Delta \psi + k^2 \psi = 0
\]  

is expressed as

\[
\psi(r) = u_0(r) = \int c(s) e^{ik(r \cdot s)} d\omega,
\]

where \( r \) is the position vector, \( s \) a unit vector, and \( d\omega \) an infinitesimal solid angle around \( s \). \( u_0(r) \) can be shown, by means of partial integration, to tend for \( r \to \infty \) to the asymptotic spherical wave

\[
u_0(r) = \frac{4\pi}{k} \left| c \left( \frac{r}{r} \right) \right| \frac{1}{r} \sin k(r + \gamma),
\]

with \( \gamma \) being defined by \( c = |c| e^{ik\gamma} \). Therefore, "the intensity as a function of the direction \( s = r / r \) determines the frequency of occurrence of a particle coming into the solid angle element \( d\omega \) with \( s \) as its axis" to be

\[
\Phi_0 d\omega = |c(s)|^2 d\omega.
\]

Schrödinger's differential equation for a three-dimensional motion under the influence of potential energy \( U(r) \),

\[
\Delta \psi + k^2 \psi = V \psi, \quad V(r) = \left( \frac{8\pi^2 \mu}{\hbar^2} \right) U(r),
\]
can be solved by the iterative method mentioned above if $V(r) \to \text{const.}/r^2$ as $r \to \infty$. Namely, in terms of the solution to the inhomogeneous differential equation

$$\Delta u_n + k^2 u_n = V u_{n-1},$$

the series

$$\psi(r) = \sum_{n=0}^{\infty} u_n(r)$$

is the convergent solution of Eq. (161). $u_n(r)$ is given by

$$u_n(r) = -\frac{1}{4\pi} \int V(r') u_{n-1}(r') \frac{e^{-ik|r-r'|}}{|r-r'|} dr'.$$

Putting $u_n \to u_n^\infty$ as $r \to \infty$ and

$$f_{n-1}(ks) = (2\pi)^{-3} \int V(r') u_{n-1}(r') e^{-ik(s\cdot r')} dr',$$

one gets

$$u_n^\infty(r) = 2\pi^2 f_{n-1}(-ks) \frac{e^{-ikr}}{r},$$

with the use of partial integration.

The comparison of this with Eqs. (159) and (160) gives therefore "the probability that an electron is deflected into a solid angle element $d\omega$ around the mean direction $s$" as

$$\Phi d\omega = \pi^2 k^2 \left| \sum_{n=0}^{\infty} f_n(-ks) \right|^2 d\omega.$$  

Taking particularly the first approximation, one has

$$\Phi d\omega = \pi^2 k^2 |f_0(-ks)|^2 d\omega.$$  

This is the formula for the probability of scattering, which is called the Born approximation nowadays. Born also treats inelastic scatterings on the supposition that the potential energy of the electron depends on the internal coordinates of the atom too. The mathematical procedure used by him is, as in

\footnote{As Born's Eq. (166) is multiplied with the time factor $\exp(ikt)$, $\exp(-ikr)/r$ corresponds to the outgoing spherical wave.}
the case of elastic scatterings mentioned above, almost the same as what is included in present-day textbooks on quantum mechanics.

In this way Born related Schrödinger’s wave function $\psi$ to the particle nature of electron by analyzing the problem of scattering, and arrived at the idea, which is equivalent to one of the most fundamental concepts in quantum mechanics, that $|\psi|^2$ should represent the probability of finding a particle.\(^{59}\)

However, the time-dependent wave equation used by Born was, as is described above, of second order in the time derivative, and was the same as that used by Schrödinger in his first and second papers and in the paper on the relationship, which were cited in Born’s paper under consideration. The wave equation of first order in the time derivative presented by Schrödinger in his fourth paper was not yet used in this Born’s work.\(^{60}\)

If $\psi$ is to give the probability and to obey a partial differential equation of second order in the time derivative, the probability at an arbitrary time could not be determined unless both the initial values of the probability and the rate of its time variation are given. This point did not explicitly appear in Born’s analysis of the problem of scattering, because the “aperiodic system” considered in it referred to a system that had continuous energy eigenvalues and did not correspond to a classical periodic system making a rotational orbital motion. If, on the other hand, $\psi$ obeys a partial differential equation of first order in the time derivative, the probability at any time is determined when its initial value is given.\(^{61}\) Born, on the basis of the partial differential equation of first order in the time derivative which was proposed by Schrödinger together with the interpretation of $|\psi|^2$ as the “weight function”, came later to establish the interpretation of $|\psi|^2$ as the “probability”. This process will be described in the following chapter.

---

\(^{59}\) The supposition that $|c_n|^2$ should be the probability of finding the system in state $n$ is, on account of Eq. (133), equivalent to the supposition that $|\lambda(q)|^2 \, dq$ should be the probability of finding the system in the coordinate interval $dq$. This was pointed out by Pauli (cf. Sec. 2, §6, Chap. 3 of the present volume).

\(^{60}\) Born’s paper under consideration and Schrödinger’s fourth paper were received at nearly the same time, as we have mentioned before.

\(^{61}\) In Schrödinger’s derivation of the dispersion formula of light, which we have mentioned in §4 of the last section, Schrödinger solved the time-dependent wave equation of first order in the time derivative, Eq. (116), by putting the wave function in the form of Eq. (117), as in the case of a forced oscillation, so that no consideration was given of the initial condition.
This page is intentionally left blank
Chapter 3

The Establishment of Quantum Mechanics

1. Interpenetration between Matrix- and Wave-Mechanics

§1. Matrix- and wave-mechanics for many-body systems

As we have analyzed in detail in Chap. 1 of the present volume, Heisenberg's idea of getting combination rules for the quantum-theoretical amplitudes which belong to an atom corresponding to its various transitions was formulated by Born and Jordan as matrix mechanics, with its further development made by Born, Heisenberg and Jordan, and was also formulated by Dirac as mechanics modified by quantum algebra, which was mathematically equivalent to matrix mechanics. In Chap. 2, we have then followed the process in which Schrödinger formulated de Broglie's idea of phase wave into wave mechanics.

In this process Schrödinger showed, in his paper on the relationship1) received on Mar. 18, 1926, that matrix elements in matrix mechanics could in general be given in terms of the eigenfunctions of the wave equation, as we have described in Sec. 3 of the last chapter. Following this result, arrived at by Schrödinger, Heisenberg and Dirac independently applied, from their own respective standpoints, matrix and wave mechanics, in comparison with each other, to many-body systems consisting of identical particles.

Heisenberg says, in his paper "Many-body Problem and Resonance in the Quantum Mechanics"2) received on Jun. 11, 1926, that "quantum mechanics has been until now applied only to systems which consist of one moving mass point", on account of "the mathematical difficulties which stood in the way of

1) E. Schrödinger, Ref. 27) in Sec. 3, §1, Chap. 2 of the present volume.
calculating individual elements of amplitude”. Here, by “elements of amplitude” Heisenberg means matrix elements and by “quantum mechanics”, matrix mechanics. Heisenberg considers, however, that “recently in this problem an extraordinary progress has been achieved by the significant investigations, in which Schrödinger has discovered a new, mathematically much more comfortable access to the realm of quantum mechanics”.

Nevertheless, Heisenberg argues that “because of the mathematical equivalence of Schrödinger’s method to quantum mechanics, the question about the events which give rise to the equations could temporarily be considered as a problem of appropriateness concerned with our intuition, but only when we do not attempt to extend the basis of these theories, on the ground of the intuitional image once chosen”, and stresses the dominance of matrix mechanics over wave mechanics in many-body problems, as follows.

“As far as can be seen” by Heisenberg, “Schrödinger’s method does not represents a consistent wave theory of matter in the sense of de Broglie”. According to Heisenberg, the extensive application of the wave theory to the \( f \)-dimensional space and the calculation of wave velocity from the mutual potential energy of the particles, as in Schrödinger’s work, mean “borrowing from the concepts of the corpuscular theory”. Furthermore, Heisenberg says that “even if a consistent wave theory of matter in the ordinary three-dimensional space could be developed in accordance with the program of de Broglie and Einstein, an exhaustive description of atomic processes in our ordinary space-time concepts would hardly be obtained in this way”.

Namely, Heisenberg thinks that “just in view of the close-fitting analogy between light and matter, which becomes more and more evident, one might believe that such a wave theory of matter would hardly enable a complete description of our atomic experiences, just as the wave theory of light does not enable any interpretation of optical experiences”. “In view of this analogy”, to Heisenberg “it seems to be one of the most important sides of quantum mechanics, that quantum mechanics is based on the corpuscular idea”.

As we have already mentioned in Chaps. 1 and 2 of the present volume, matrix mechanics on one hand was based on Heisenberg’s phenomenological view that there should be “aggregate” of quantum amplitudes corresponding to a physical quantity such as the classical amplitude of an atomic electron, and wave mechanics on the other hand stood on Schrödinger’s substantialistic view.

\(^3\)Heisenberg cites here Schrödinger’s first and second papers and the paper on the relationship.
to treat the de Broglie wave as "an existence in an intuitional and concrete form". The conflict between matrix mechanics and wave mechanics thus meant that between a phenomenological theory and a substantialistic theory, and that between a corpuscular theory and an undulatory theory.

Heisenberg’s theory was an extension of Bohr’s correspondence principle, in which the frequency and amplitude of a quantum-theoretical transition were supposed to correspond to those of a Fourier component in classical theory. Schrödinger, following de Broglie's idea of phase wave, grasped the quantum condition to correspond to the theory of Fresnel zones by noting the similarity of the Fermat principle to the Hamilton principle, and developed this correspondence by means of the substance of $\psi$-wave, to show that any matrix element can be given by eigenfunctions of the wave equation. Any substantialistic theory introduces some substance, by taking its correspondence with something in a theory hitherto established as the stepping stone, and therefore possesses of itself a correspondence-theoretical aspect. It cannot be seen to be a mere absurdity, as was done by Heisenberg, that the mutual interaction energy of particles based on classical theory was used in Schrödinger’s theory.

In his second paper Schrödinger noted of his wave equation that “one is not allowed to forget that it is only just about one-body problems that the interpretation as oscillations in the actual three-dimensional space is immediately permitted”, as we have described in Sec. 2, §2 of the last chapter. In this point, the knot of the contradiction inherent in wave mechanics as opposed to matrix mechanics should be said to consist. On this account, Schrödinger reinterpreted $|\psi|^2$ as the “weight function” in the $f$-dimensional configuration space, saying at the same time that “this reinterpretation at first sight may give a shock”, in the fourth paper received ten days later than Heisenberg’s paper under consideration. And, as we have seen in Sec. 5 of the last chapter, Born opened the way, in his paper received a further four days later, to grasp $|\psi|^2$ as the probability density, “in accordance with the program of de Broglie and Einstein” if expressed in Heisenberg’s words quoted above.

Now, Heisenberg attempts in his paper under consideration to make “the quantum-mechanical treatment of systems which consist of many mass points”, that is, a matrix mechanical treatment of many-body problems, from the viewpoint that wave mechanics, which is, as a substantialistic theory in opposition to the phenomenological matrix mechanics standing on a corpuscular-theoretical ground, should be significant only as a theory of wave in the three-dimensional space in which we live. In this attempt, Heisenberg regards
the equivalence, shown by Schrödinger between the wave and matrix mechanics for systems of an arbitrary degree of freedom, phenomenologically as a convenient mathematical means to his end.

Heisenberg says, however, that “such a treatment seems to meet at first with considerable difficulties”. Namely, to him “the features of de Broglie’s theory, which lead to Einstein-Bose’s statistics, seem to possess no analogy in quantum mechanics”. As we have described in Sec. 2, §4, Chap. 5, Vol. II, Einstein attributed the cause of the mutual dependence of molecules in the Bose-Einstein statistics, in contrast to their independence in the classical statistics, to the interferential nature of the de Broglie wave. It is needless to say that interferences cannot be thought to occur in any corpuscular theory. Heisenberg says further that “additional rules, such as Pauli’s prohibition of equivalent orbits,\footnote{W. Pauli, Zeit. f. Phys. \textbf{31} (1925), 765 [same as Ref. 68) in Sec. 3, §2, Chap. 4, Vol. II]. In this paper the Pauli exclusion principle in the present-day nomenclature was presented.} have in their form no place in the mathematical scheme of quantum mechanics”. He thus tries to find something corresponding to them.

The simplest many-body problem is a system of two coupled oscillators. Each oscillator is supposed to have an identical mass \( m \) and an identical frequency \( \nu \). If their mutual interaction energy is assumed to be given by \( 4\pi^2m\lambda q_1q_2 \), in terms of their respective coordinates \( q_1 \) and \( q_2 \) and a constant \( \lambda \), “the principal oscillation” of the system is given by the coordinate transformation \( q_\pm = (q_1 \pm q_2)/\sqrt{2} \), where \( q_+ \) represents the motion in phase of the two oscillators, and \( q_- \) that in opposite phase. With \( \nu_\pm \) defined by \( \nu_\pm^2 = \nu^2 \pm \lambda \), the energy of a stationary state of the system is given by \( h\nu_+(n_++1/2) + h\nu_-(n_-+1/2) \), where \( n_+ \) and \( n_- \) are integers, so that each spectral term of the system is specified by a pair of quantum numbers \( (n_+n_-) \).

If the two oscillators represent charged particles, the electric dipole moment is proportional to \( q_+ \). “In the first approximation there occur therefore only transitions of \( n_1 \) by 1”. Heisenberg uses the selection rule to state this in the approximation of electric dipole radiation. As we have seen in Sec. 2, §5, Chap. 1 of the present volume, the selection rule which had played an important role in the old quantum mechanics was derived from matrix mechanics by Born-Heisenberg-Jordan. In higher approximations, the radiation due to the quadruple and higher poles becomes effective. These higher terms are given in terms of \( q_1, q_2 \) and their time derivatives. When expressed in \( q_+ \) and \( q_- \), they include as factors an even number of \( q_- \) and its time derivative. Heisenberg says
therefore that “also, by taking account of all higher approximations, \( q_- \) can always change in radiation only by an even number”.

The spectral terms thus make, for each value of \( n_- \), a series \((0n_-), (1n_-), (2n_-), \ldots\), and no combination occurs between a series of even \( n_- \) and that of odd \( n_- \). “The term scheme can therefore be split into two partial systems”. Heisenberg asks now the question “Are both the partial systems realized in the nature, or only one of them?"

Of this question he says that “the indetermination of the quantum mechanical answer seems to be the most essential result of the study”. He however states that “this gives just as much freedom such that the demands of Bose-Einstein’s counting and Pauli’s prohibition of equivalent orbits can be fitted naturally into the system of quantum mechanics”. Wishing to “show that coupled systems always stand in quantum mechanics similar to the example explained above”, Heisenberg treats then the “resonance” between two atoms.

§2. Heisenberg’s analysis of the “resonance” between atoms

In classical mechanics, two periodically oscillating systems can enter into actual resonance only when their frequencies are independent of the respective energies and are approximately equal to each other. Heisenberg sees, “in accord with the general experiences”, that “in quantum mechanics two atomic systems always enter into resonance when the absorption frequency of one system agrees with the emission frequency of the other, or vice versa”, and concludes that “because of the character of linearity inherent in the quantum-mechanical equations, the resonance in quantum mechanics is a phenomenon more general than in classical mechanics”.

In order to study this problem in detail, he takes two identical atoms \( a \) and \( b \) of \( f \) degrees of freedom. They are supposed to be coupled to each other through the interaction energy \( H' \). The energies of their stationary states are denoted by \( H^a_n \) and \( H^b_m \) respectively, with \( n \) and \( m \) representing the sets of the respective quantum numbers. A stationary states of the whole system is specified by a pair \( nm \) of the quantum numbers \( n \) and \( m \). Without taking account of the mutual interaction, its energy is given by

\[
H_{nm} = H^a_n + H^b_m. \tag{1}
\]

“The whole system possesses now the degeneracy characteristic of the resonance”. In other words, because of the relation \( H_{nm} = H_{mn} \) every energy
eigenvalue is a doublet except for the case \( n = m \), so that "resonance always occurs when the two systems were initially not in the same state". "In the system disturbed by the interaction the degeneracy is canceled", and "there results a secular beat, in which the energy pulsates to and from between the two partial systems". "Formally the connection is as follows". The energy of the disturbed system \( W' \) is given "in the first approximation by the time average of \( H' \) over the undisturbed motion". "This average value will still include in general terms corresponding to the transitions, in which the systems \( a \) and \( b \) interchange the places". "One must hence carry out a canonical transformation so that \( W' \) becomes a diagonal matrix".\(^5\)

Heisenberg expresses this canonical transformation as

\[
W' = S^{-1} \bar{H}' S, \quad q' = S^{-1} q S,
\]

where \( \bar{H}' \) contains only diagonal elements and those elements corresponding to the transitions between states of the same energy. Namely, decomposing \( SW' = \bar{H}' S \), one gets the following simultaneous equations,

\[
\begin{cases}
(W' - H'_{nm,nm})S_{nm,nm} - H'_{nm,mn}S_{mn,nm} = 0 \\
-H'_{mn,nm}S_{nm,nm} + (W' - H'_{mn,mn})S_{mn,nm} = 0
\end{cases}.
\]

From the symmetry of \( H' \) with respect to the atoms \( a \) and \( b \), one has

\[
H'_{nm,nm} = H'_{mn,mn}, \quad H'_{nm,mn} = H'_{mn,nm}.
\]

Equation (3) has two sets of solutions given by

\[
\begin{cases}
W'_{nm} = H'_{nm,nm} + H'_{nm,mn} \\
S_{nm,nm} = S_{mn,nm} = 1/\sqrt{2}
\end{cases},
\]

\[
\begin{cases}
W'_{mn} = H'_{nm,nm} - H'_{nm,mn} \\
S_{nm,nm} = -S_{mn,nm} = 1/\sqrt{2}
\end{cases}.
\]

"The term spectrum splits again into two series of terms, which can in no way combine with each other". Heisenberg says that "this is the decisive result". "Obviously the same relationship rules here as in the special example".

given above. The non-occurrence of combinations between the two kinds of series of terms is here again related to "the equality of the coupled systems". Furthermore, "the quantum-mechanical solution has again the above stressed degree of uncertainty" of the one to be realized out of the two kinds of series of terms".

Heisenberg then attempts to explain by means of wave mechanics the result thus obtained, saying that "it may be interesting to study the problem of resonance just described from the standpoint of Schrödinger's method". It is assumed that "Schrödinger's eigenfunction" $\varphi_n^a$ belonging to the partial state $a$ with energy $H_n^a$ is known, and this is similarly the case of the partial system $b$. The eigenfunction belonging to the state of the whole system with energy $H_{nm}$ is given by $\varphi_n^a \varphi_m^b$. The eigenfunctions of the perturbed system corresponding to Eqs. (5a) and (5b) are obtained by applying the matrix $S$ to the needed linear transformation of the eigenfunctions. They are given by

$$\begin{align*}
(\varphi_n^a \varphi_m^b + \varphi_m^a \varphi_n^b)/\sqrt{2} & \quad \text{for } W_{nm}' \quad \text{(6a)} \\
(\varphi_n^a \varphi_m^b + \varphi_m^a \varphi_n^b)/\sqrt{2} & \quad \text{for } W_{mn}' \quad \text{(6b)}
\end{align*}$$

Heisenberg takes an integral of the type

$$\int F\left(\varphi_n^a \varphi_m^b + \varphi_m^a \varphi_n^b\right)\overline{\varphi_n^a \varphi_m^b} - \overline{\varphi_m^a \varphi_n^b}dq_1^a \cdot \cdot \cdot dq_{2n}^a,$$

where $F$ is "a function or an operator symmetric in $a$ and $b''$, $\overline{\varphi_n^a}$ etc. being the complex conjugate of $\varphi_n^a$ etc. This integral "must be zero", because "it changes the sign when the systems $a$ and $b$ are exchanged, but at the same time still keeps the same value". He says that "the non-existence of the intercombination between the two partial systems follows simply from this". That is to say, Heisenberg sees in this simplicity the "mathematically much more comfortable" nature of Schrödinger's wave function.

So, getting the result given in Eqs. (5a) and (5b) and by the method of first order perturbation for a degenerate system based on matrix mechanics,\(^6\) Heisenberg arrives, under the guidance of this result, at the idea of a wave function symmetric or antisymmetric with respect to the exchange between two identical partial systems, such as given in Eqs. (6a) or (6b). The symmetry

\(^6\)In Heisenberg's Eq. (2) or (3), only those elements of the matrix $H'$ that are relevant to the two degenerate states $nm$ and $mn$ are taken into account. This is the characteristics of the method of first order perturbation for a degenerate system.
property of this kind for the exchange between two identical systems was not hitherto considered in Schrödinger’s wave mechanics.

Heisenberg then generalizes this idea to a system consisting of \( n \) identical partial systems. The \( n! \) states obtained by exchanging the \( n \) identical partial systems with each other have one and the same energy if no interaction works among the partial systems. By the interaction the degeneracy is canceled, and “the system of spectral terms splits into \( n! \) partial systems on account of resonance”. “Among them there is a system which contains no equivalent orbit and cannot combine with the other systems”. Heisenberg shows this “shortly” as follows.

If the interaction energy \( H' \) consists of the sum of the interactions between a pair of partial mechanical systems, “\( H' \) contains in first approximation only such terms that correspond at most to the transitions between two partial mechanical systems”. If a state of the whole system is denoted by the set of the “quantum numbers” \( (m_1, m_2, \ldots, m_n) \) of the partial systems, and the \( n(n-1)/2 \) nondiagonal matrix elements of \( H' \) relevant to transitions between two states realized by exchanging a pair of the quantum numbers, are denoted by \( H'_\beta, \ldots, H'_\gamma, H'_\nu \), while its \( n \) diagonal elements corresponding to no transition are denoted by \( H'_\alpha \), one has a set of simultaneous equations to determine the canonical transformation, which generalizes Eq. (3). The determinant of its coefficients is

\[
\begin{vmatrix}
W' - H'_\alpha & -H'_\beta & -H'_\gamma & \cdots & -H'_\nu & 0 & \cdots & 0 \\
-H'_\delta & W' - H'_\alpha & 0 & \cdots & 0 & -H'_\epsilon & \cdots & \cdots \\
-H'_\eta & 0 & W' - H'_\alpha & \cdots & 0 & -H'_\zeta & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\end{vmatrix}
\]

(8)

One of the solutions obtained by putting this determinant equal to zero is given by

\[
W' = H'_\alpha - H'_\beta - H'_\gamma \cdots - H'_\nu, \quad S_k = \frac{(-1)^{\delta_k}}{\sqrt{n!}},
\]

(9)

where \( k \) denotes the state \( (m^k_\alpha, m^k_\beta, \ldots, m^k_\nu) \) which results from \( \delta_k \) times of permutations between two quantum numbers starting from the original state \( (m_1, m_2, \ldots, m_n) \), and \( S_k \) is the element of the matrix \( S \) between the state \( k \) and the original one. “Schrödinger’s eigenfunction belonging to this solution”
is given by

\[ \varphi = \left( \frac{1}{\sqrt{n!}} \right) \sum (-1)^{\delta_k} \varphi_1(m_{\delta_1}^k) \varphi_1(m_{\delta_2}^k) \cdots \varphi_n(m_{\delta_n}^k). \] (10)

This \( \varphi \) does not contain any combination of partial systems in identical state. If \( \psi \) is an eigenfunction that contains at least one such combination, one has for a function or operator \( F \) symmetric in the partial systems

\[ \int F\psi \varphi d\Omega = 0. \] (11)

From this Heisenberg concludes that "terms of the type of Eq. (10) really build a closed system, that is, do not combine with any other term".

Equation (10) is identical with the determinant wave function, so-called nowadays, for a system of many bodies obeying the Fermi-Dirac statistics. This does not mean, however, that the Fermi-Dirac statistics is due to Heisenberg's paper under consideration. Heisenberg considers rather that the existence of such a solution among the \( n! \) solutions "corresponds to the reduction of the statistical weight in Bose-Einstein's counting from \( n! \) to 1". He concludes however that "the formulation given here of this counting goes beyond Bose-Einstein's in this respect, because it prescribes the selection of a fully determined system from the \( n! \) solutions, namely such a system that contains no equivalent system and therefore is in accord with Pauli's prohibition".

Thus, Heisenberg says that "a reasoning of the fact, that just only this single system among all the possible quantum-mechanical solutions occurs, cannot be given only from quantum-mechanical calculations". To him it seems, however, to be "an important result" of his study "that Pauli's prohibition and Einstein's statistics have the same ground, and that they do not contradict with quantum mechanics".

In this way, Heisenberg treated the mutual dependence among atoms, on the analogy of the inter-molecular dependence in the Bose-Einstein statistics, and found that the result "follows simply" from the permutational symmetry of the wave function. However, Heisenberg did not come to propose the Fermi-Dirac statistics by pushing out the permutational symmetry of the wave function, because he considered phenomenologically the equivalence of wave mechanics with matrix mechanics as a "mathematically much more comfortable" means.

---

7) Eq. (10) is Eq. (20) in the original paper.
Differently from Heisenberg, Dirac clarified the difficulty in matrix mechanics for a system of identical particles, as we shall see in the following subsection, and solved it by pushing out the permutational symmetry of the wave functions. Dirac arrived, on the basis of his result, at the Fermi-Dirac statistics as well as the Bose-Einstein statistics.

§3. Dirac's analysis of a system of identical particles

Dirac treats, in his paper "On the Theory of Quantum Mechanics" received on Aug. 26, 1926, "a system containing several similar particles, such as an atom with several electrons". Taking "for definiteness" an atom with two electrons, Dirac denotes with \((mn)\) the state of the atom in which one electron is "in an orbit labeled \(m\)”, and the other "in the orbit \(n\)".

Dirac asks now the question "whether the two states \((mn)\) and \((nm)\), which are physically indistinguishable as they differ only by the interchange of the two electrons, are to be counted as two different states or as only one state”, that is, whether they give rise to "two rows and columns in the matrices or to one only”.

If the first alternative is right, one would be able from "Heisenberg's matrix mechanics" to calculate "separately" the intensities due to the two transitions \((mn) \to (m'n')\) and \((mn) \to (n'm')\). "The two transitions are, however, physically indistinguishable, and only the sum of the intensities for the two together could be determined experimentally”. "Hence, in order to keep the essential characteristic of the theory that it shall enable one to calculate only observable quantities, one must adopt the second alternative that \((mn)\) and \((nm)\) count as only one state”.

"This alternative, though, also leads to difficulties”. From the "symmetry" between the two electrons, it is required that the amplitude associated with the transition \((mn) \to (m'n')\) of the coordinate \(x_1\) in the \(x\)-direction of the first electron, should be equal to the amplitude associated with the transition \((nm) \to (n'm)\) of the corresponding coordinate \(x_2\) of the other electron. That is, in the form of matrix element one should have the relation.

\[
x_1(mn;m'n') = x_2(nm;n'm').
\]  
(12)

If, on the other hand, \((mn)\) and \((nm)\) are both counted as defining the same row and column of the matrices, Eq. (12) shows that each element of the matrix

$x_1$ is equal to the corresponding element of the matrix $x_2$, so that there should hold “the matrix equation”

$$x_1 = x_2.$$  \hspace{1cm} (13)

Dirac says that “this relation is obviously impossible, as, amongst other things, it is inconsistent with the quantum conditions”. He means by “the quantum conditions” the commutation relations for canonical coordinates and momenta, from which the relation $x_1 x_2 = x_2 x_1$ follows for the two independent matrices $x_1$ and $x_2$.

Dirac concludes that “one consequence of these considerations is that the theory of uniformising variables can no longer apply”. “The theory of uniformising variables” is the one due to Dirac in which the method of multiple Fourier expansion for a multiple-period motion in classical mechanics is converted into quantum mechanical one by means of the “q-algebra”, as we have seen in Sec. 3, §2, Chap. 1 of the present volume. Dirac concludes that “as corresponding to any transition $(mn) \rightarrow (m'n')$ there is a term $e^{i(\alpha w)}$ in the Fourier expansion, we should require there to be a unique state, say $(m''n'')$, such that the same term $e^{i(\alpha w)}$ corresponds to the transition $(m'n') \rightarrow (m''n'')$, and $e^{2i(\alpha w)}$ corresponds to $(mn) \rightarrow (m''n'')$”. If the $m$’s and $n$’s are quantum numbers, we have therefore to take

$$m'' - m' = m' - m, \quad n'' - n' = n' - n.$$  

“Since, however, the state $(m'n')$ may equally well be called the state $(n'm')$, we may equally well take

$$m'' - n' = n' - m, \quad n'' - m' = m' - n,$$

which would give a different state $(m''n'')$”. “There is thus no unique state $(m''n'')$ that the theory of uniformising variable demands”.

After having clarified the difficulty inherent in matrix mechanics applied to a system of identical particles, Dirac attempts to solve this difficulty by the symmetry of Schrödinger’s wave function for the permutation of particles. If we neglect for the moment the interaction between the two electrons and denote the eigenfunction for a single electron in the orbit $m$ by $\psi_m(x, y, z, t)$, the eigenfunction for the whole atom in the state $(mn)$ is given by$^9$

$$\psi_m(x_1, y_1, z_1, t)\psi_n(x_2, y_2, z_2, t),$$  \hspace{1cm} (14)

$^9$For the time variable $t$ in $\psi_m$ and $\psi_n$, see §5 of the present section.
The Establishment of Quantum Mechanics

which will be abbreviated as

$$\psi_m(1)\psi_n(2).$$

If, then, we count the \((mn)\) and \((nm)\) states as identical, the eigenfunction \(\psi_m(2)\psi_n(1)\) corresponds to the same state of the atom. However, because “two independent eigenfunctions must give rise to two rows and columns in the matrices”, “we must find a set of eigenfunctions of the form

$$\psi_{mn} = a_{mn}\psi_m(1)\psi_n(2) + b_{mn}\psi_m(2)\psi_n(1),$$

where the \(a_{mn}\)'s and \(b_{mn}\)'s are constants”, in order “to have only one row and column in the matrices corresponding to both \((mn)\) and \((nm)\)”. This set of eigenfunctions “must contain only one \(\psi_{mn}\) corresponding to both \((mn)\) and \((nm)\), and must be sufficient to enable one to obtain the matrix representing any symmetrical function \(A\) of the two electrons”. This means that the \(\psi_{mn}\)'s must be chosen so that \(A\) times any chosen \(\psi_{mn}\) can be expanded in terms of the chosen \(\psi_{mn}\)'s in the form

$$A\psi_{mn} = \sum_{m'n'} \psi_{m'n'}A_{m'n',mn},$$

where \(A_{m'n',mn}\)'s are constants or functions of the time only.

Dirac says that “there are two ways of choosing the set of \(\psi_{mn}\)'s to satisfy the conditions”. In one of them, we may take \(a_{mn} = b_{mn}\), which “makes each \(\psi_{mn} \) a symmetrical function of the two electrons”, so that the left-hand side of Eq. (16) becomes symmetrical and only symmetrical eigenfunctions will be required for its expansion. In the other way, we may take \(a_{mn} = -b_{mn}\), which “makes \(\psi_{mn} \) antisymmetrical”, so that the left-hand side of Eq. (16) becomes antisymmetrical and only antisymmetrical eigenfunctions will be required for its expansion. He considers that, though “the theory at present is incapable of deciding which solution is the correct one”, “we are able to get complete solutions of the problem which make use of less than the total number of possible eigenfunctions at the expense of being able to represent only symmetrical functions of the two electrons by matrices”.

“These results may evidently be extended to any numbers of electrons”. For \(r\) non-interacting electrons, the symmetrical eigenfunctions are

$$\sum_{\alpha_1\alpha_2\cdots\alpha_r} \psi_{n_1}(\alpha_1)\psi_{n_2}(\alpha_2)\cdots\psi_{n_r}(\alpha_r),$$

(17)
where $\alpha_1 \alpha_2 \cdots \alpha_r$ are any permutation of the integers $1, 2, \ldots, r$, while the antisymmetrical one "may be written in the determinantal form"

\[
\begin{vmatrix}
\psi_{n_1}(1) & \psi_{n_1}(2) & \cdots & \psi_{n_1}(r) \\
\psi_{n_2}(1) & \psi_{n_2}(2) & \cdots & \psi_{n_2}(r) \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{n_r}(1) & \psi_{n_r}(2) & \cdots & \psi_{n_r}(r)
\end{vmatrix}.
\]

(18)

Dirac says that if there is interaction, "there will be symmetrical and antisymmetrical eigenfunctions, although they can no longer be put in these simple forms", and stresses that "in any case the symmetrical ones alone or the antisymmetrical ones alone give a complete solution of the problem". He points out clearly that "in the solution of the problem with antisymmetrical eigenfunctions there can be no stationary states with two or more electrons in the same orbit, which is just Pauli's exclusion principle", but "the solution with symmetrical eigenfunctions cannot be the correct one for the problem of electrons in an atom".

In this way, Dirac solves the difficulty over identical particles in matrix mechanics of many-body problems by introducing antisymmetry and symmetry of the wave functions for the exchange of two identical particles. Dirac cites Schrödinger's theory by saying in a footnote "See various papers in the 'Ann. d. Phys.,' beginning with vol. 79, p. 361 (1926)" and states that "the mathematical equivalence of the theories is thus established" by Schrödinger. This establishment was done in Schrödinger's "paper on the relationship", which we have mentioned in Sec. 3, Chap. 2 of the present volume. As we have seen in §3 of the same section, Schrödinger analyzed in this paper the difference between matrix mechanics and wave mechanics, and considered the former as a phenomenological theory and the latter as a substantialistic theory, if expressed in the terminology of our three-stage theory.

Dirac's expression quoted above "the essential characteristic of Heisenberg's theory that it shall enable one to calculate only observable quantities" indicates also the phenomenological character of matrix mechanics. Noting that

10) W. Pauli, Zeit. f. Phys. 31 (1925), 765 [the same as Ref. 68] in Sec. 3, §2, Chap. 4, Vol. II.

11) It is not shown explicitly up to which paper of Schrödinger the term "various papers" means to include. Dirac cites, in another paper received on Dec. 2, 1926, which we shall consider in Sec. 2, §4 of the present chapter, Schrödinger's fourth paper.
the method of matrix mechanics does not apply as it is to a system of identical particles, Dirac adopted the result achieved by wave mechanics that phenomenological matrix elements can be calculated if substantialistic wave functions are known. On the other hand, in spite of various trials Schrödinger did not come to clarify, from the standpoint of wave mechanics, the connection between a transition probability and the corresponding matrix element, while this connection was premised phenomenologically in matrix mechanics. Dirac determined the matrix elements to give transition probabilities of an atom having two or more electrons by imposing the symmetry requirement on the wave function. That is, Dirac was going to the synthesis of the phenomenological and substantialistic theories.

§4. Bose-Einstein statistics and Fermi-Dirac statistics

In the paper under consideration, Dirac applies then the result mentioned above that the wave function of a system of identical particles should be symmetric or antisymmetric for the exchange of two particles, to an assembly of gas molecules. He shows that symmetric wave functions give rise to the Bose-Einstein statistics, and antisymmetric wave functions to the Fermi-Dirac statistics.

Dirac states first that "the wave equation for a single molecule of rest mass $m$ moving in free space is

$$\left\{ p_x^2 + p_y^2 + p_z^2 - \frac{W^2}{c^2} + m^2c^2 \right\} \psi = 0, \quad (19)$$

$$\left\{ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{m^2c^2}{\hbar^2} \right\} \psi = 0''. \quad (20)$$

These equations$^{12}$ are derived from the considerations which Dirac develops in the early parts of the paper. We shall treat them in the following subsection §5. The solution of Eq. (20) has the form

$$\psi_{\alpha_1\alpha_2\alpha_3} = \exp \left[ \frac{i}{\hbar} (\alpha_1x + \alpha_2y + \alpha_3z - Et) \right], \quad (21)$$

$^{12}$Equations (19) and (20) are not numbered in the original paper. By $\hbar$ Dirac denotes $(2\pi)^{-1}$ times the usual Planck constant. In the present section we use $\hbar$ instead of Dirac's $\hbar$ to avoid confusion.
where \( \alpha_1, \alpha_2, \alpha_3 \) and the energy \( E \) are constants satisfying

\[
\alpha_1^2 + \alpha_2^2 + \alpha_3^2 - E^2 + m^2c^2 = 0.
\] (22)

This eigenfunction represents an atom having the momentum components \( \alpha_1, \alpha_2, \alpha_3 \) and the energy \( E \).

One must now look into "some restriction on the possible eigenfunctions due to the presence of boundary walls". On the supposition that each molecule is confined between two boundaries at \( x = 0 \) and \( x = 2\pi \), Dirac assumes that, "as a natural generalization" of the usual methods, "only those functions of \( x \) that are defined only for \( 0 < x < 2\pi \) have a physical meaning and must be capable of being represented by matrices". These functions \( f(x) \) can always be expanded as Fourier series. Therefore, if one chooses from the eigenfunctions given in Eq. (21) those for which \( \alpha_1/h \) is an integer, \( f(x) \) times any chosen eigenfunction can be expanded as a series in the chosen eigenfunctions, and hence \( f(x) \) can be represented by a matrix [cf. Eq. (16)]. This holds true also for \( \alpha_1/h = n + \varepsilon \), where \( n \) is an integer and \( \varepsilon \) is any real number. Though the theory is "incapable of deciding" the values of \( \varepsilon \), "for statistical problems they all lead to the same results".

If \( y \) and \( z \) are also bounded between 0 and \( 2\pi \), "the number of waves associated with the molecules" whose energies lie between \( E \) and \( E + dE \) is given by

\[
\left( \frac{4\pi}{c^3h^3} \right) (E^2 - m^2c^4)^{-1/2} E dE.
\] (23)

This value is in agreement with the value obtained on "the ordinary assumption that the wave function vanishes at the boundary". For nonrelativistic cases, Eq. (23) reduces to

\[
\left( \frac{2\pi}{h^3} \right) (2m)^{3/2} E_1^{1/2} dE_1,
\] \[ E_1 = E - mc^2. \] (24)

If the volume of gas is, more generally, \( V \), this expression must be multiplied by \( V/(2\pi)^3 \).

On the assumption of no interaction between molecules, "the eigenfunctions for the assembly of molecules" are either the symmetrical eigenfunctions or the antisymmetrical ones obtained from the eigenfunctions for separate molecules. Dirac makes now "the new assumption that all stationary states of the assembly have the same \( a \ priori \) probability". Consequently, "if we adopt the
solution of the problem that involves symmetrical eigenfunctions, we should find that all values for the number of molecules associated with any wave have the same \textit{a priori} probability, which gives just the Einstein-Bose statistical mechanics”.

On the other hand, “we should obtain a different statistical mechanics if we adopt the solution with antisymmetrical eigenfunctions, as we should then have 0 or 1 molecule associated with each wave”. Since it is known that “the Einstein-Bose statistical mechanics” leads to the Planck law of black-body radiation, “the solution with symmetrical eigenfunctions must be the correct one when applied to light quanta”. Dirac says that “the solution with antisymmetrical eigenfunctions, though, is probably the correct one for gas molecules”. This is because he considers that “it is known to be the correct one for electrons in an atom, and one would expect molecules to resemble electrons more closely than light-quanta”.

As we have already seen in Sec. 2, §4, Chap. 5, Vol. II, Einstein applied Bose’s statistics to monoatomic ideal gases, and revised it to get what is called the Bose-Einstein statistics nowadays. Dirac showed that there could be a new statistics different from the Bose-Einstein statistics. Though at that time the theory was “incapable of deciding” which solution would be the correct one, the symmetrical or the antisymmetrical, as is seen in Dirac’s words quoted above, it seemed to Dirac that molecules would obey this new statistics different from the Bose-Einstein statistics. The theory that determines the statistics to be obeyed by identical particles was later obtained from an understanding of the connection between the spin of particles and statistics.

Now, in the paper under consideration, Dirac works out “the equation of state of the gas” with the use of a statistics different from that of Bose-Einstein. He divides the waves into a number of sets such that the waves in each set can be “associated with” molecules of about the same energy. If $A_s$ is the number of possible waves in the $s$th set, of which $N_s$ are “associated with” molecules, the distribution probability $w$ of $N_s$ is

$$w = \prod_s \frac{A_s!}{N_s!(A_s - N_s)!}.$$  \hfill (25)

The entropy $S$ of the gas is given by $S = k \log w$. By maximizing $S$ under the condition that the total number of molecules $N = \sum_s N_s$ and the total energy $E = \sum_s N_s E_s$ remain unaltered, where $E_s$ is the kinetic energy of a molecule
associated with a wave in the $s$th set, he gets

$$N_s = \frac{A_s}{\exp(\alpha + E_s/kT) + 1}.$$  

(26)

The application of this equation and Eq. (24) to $A_s$ with $E_1$ replaced by $E_s$ gives $N = \sum_s N_s$ and $E = \sum_s N_s E_s$ in the form of integration with respect to $E_s$, and the pressure $P$ is known from the formula $PV = (2/3)E$.

The distribution formula (26) was also obtained by Fermi in his paper “On the Quantization of the ideal mono-atomic Atom”, received on Mar. 24, 1926, about five months before Dirac’s paper under consideration, as we have mentioned in Footnote 46) in Sec. 2, §4, Chap. 5, Vol. II, though Fermi’s paper was not cited in Dirac’s paper. The statistics just described above thus became known later as the Fermi-Dirac statistics.

In this paper, Fermi takes account of Pauli’s exclusion principle and assumes that “at most one molecule with preset quantum numbers can exist in our gas”. He considers molecules in a suitable external force field instead of walls enclosing them. Saying that “the result does not depend on the choice of the force field”, he treats the force field of an isotropic harmonic oscillator. The energy $\varepsilon$ of a molecule in this force field is given by

$$\varepsilon = h\nu(s_1 + s_2 + s_3) = hv s,$$  

(27)

in terms of the frequency $\nu$ of the oscillator and the quantum numbers $s_1$, $s_2$, $s_3$ of the molecule. Molecular states with the same value of $s = s_1 + s_2 + s_3$ are of the same energy.

Since at most one molecule can exist in every state, the number of molecules of energy $hv s$ is at most $Q_s = (s+1)(s+2)/2$. If $N_s$ is the number of molecules in a gas that has energy $hv s$, one has the relation $N_s \leq Q_s$. Fermi’s $Q_s$ and $N_s$ correspond to Dirac’s $A_s$ and $N_s$ respectively. Fermi expresses the distribution probability of $N_s$ as

$$P = \prod_s \left(\frac{Q_s}{N_s}\right),$$  

(28)

with the use of binomial coefficients. From $N_s$ and the classical staying probability of a molecule of energy $hv s$, Fermi calculates the number density

\[E.\text{ Fermi, Zeit. f. Phys. 36 (1926), 902 [provisionally reported before in Lincei Rendiconti (6) 3 (1926), 145 in Italian].}

\[14)\left(\frac{Q_s}{N_s}\right) = \frac{Q_s!}{N_s!(Q_s-N_s)!}\]
of molecules at a given distance from the center of the harmonic oscillator, obtaining the gas pressure and the average energy of the molecule. Each of them is given as a function of the temperature and is independent of the frequency $\nu$ of the oscillator.

§5. \textit{Time as a canonical variable and the equation of motion}

Dirac says, at the beginning of his paper under consideration, that if

$$H(q_r, p_r) - W = 0,$$  \hspace{1cm} (29)

where $q_r, p_r$ are canonical variables and $W$ is the energy of the system, "then the wave equation for $\psi$ is

$$\left\{ H\left(q_r, -i\hbar \frac{\partial}{\partial q_r}\right) - W \right\} \psi = 0."$

As we have seen in the last chapter, this was shown by Schrödinger in his "paper on the relationship". Dirac considers Schrödinger's theory "from a slightly more general point of view, in which the time $t$ and its conjugate momentum $-W$ are treated from the beginning on the same footing as the other variables", and considers "the $p$'s and $W$ to be the differential operators

$$p_r = -i\hbar \frac{\partial}{\partial q_r}, \quad -W = -i\hbar \frac{\partial}{\partial t}.$$  \hspace{1cm} (30)

With this method, Eq. (20) in the last subsection is deduced immediately from the classical relativistic formula of the energy of a free particle.

A method which is substantially equivalent to this method of Dirac was as a matter of fact presented in Schrödinger's fourth paper. As we have described in Sec. 4, §3 of the last chapter, Schrödinger arrived after a series of trials, in his fourth paper received about two months before Dirac's paper under consideration,\textsuperscript{15} at the equation governing the time dependence of the wave function, that is, which is what is called the time-dependent Schrödinger equation nowadays. In the fourth paper Schrödinger proposed a relativistic wave equation, stating that "the relativistic-magnetic generalization of the basic equation" would be necessary in order to treat "Lorentz's electron".

Namely, Schrödinger wrote the relativistic Hamilton-Jacobi partial differential equation for an electron in an electromagnetic field represented by scalar

\textsuperscript{15}Cf. Footnote 11) in §3 of the present section.
and vector potentials $V$ and $A$ respectively as

$$
\left( \frac{1}{c} \frac{\partial S}{\partial t} + \frac{e}{c} V \right)^2 - \left( \frac{\partial S}{\partial x} - \frac{e}{c} A_x \right)^2 - \left( \frac{\partial S}{\partial y} - \frac{e}{c} A_y \right)^2 - \left( \frac{\partial S}{\partial t} - \frac{e}{c} A_t \right)^2 - m^2 c^2 = 0, \quad (31)
$$

where $S$ is the Hamilton principal function. In the left-hand side of this equation, he put "by a purely formal procedure"

$$
\frac{\partial S}{\partial t} = \pm \left( \frac{\hbar}{2 \pi i} \right) \frac{\partial S}{\partial t}, \quad \frac{\partial S}{\partial x} = \pm \left( \frac{\hbar}{2 \pi i} \right) \frac{\partial S}{\partial x}, \text{ etc.} \quad (32)
$$

Schrödinger then operated the differential operator so obtained on $\psi$ and equated the resultant to zero. He assumed that the equation arrived at should be "the wave equation". This wave equation becomes identical with Dirac's Eq. (20) shown above, in the case of no external field.$^{16)}$

In Schrödinger's time-dependent wave equation proposed originally, that is, Eq. (112) in Sec. 4, §3 of the last chapter, $\partial / \partial t$ was supposed to correspond to $\pm (2 \pi i / \hbar) W$. Schrödinger was unable, in the course of his trials from the standpoint of the wave mechanical monism, to choose one of the signs $\pm$. On this account, in Eq. (32) he made $p_x$ correspond to $\pm (i \hbar / 2 \pi) \partial / \partial x$, whereas in "the paper on the relationship" he took it as corresponding to $-(i \hbar / 2 \pi) \partial / \partial x$.

Next, Dirac notes the fact that if the Hamiltonian $H$ does not contain time explicitly $W$ is a constant of the system, and we can put

$$
W \psi_n = W_n \psi_n, \quad \psi_n = u_n e^{-i W_n t / \hbar}, \quad (33)
$$

in terms of a constant $W_n$ and a function $u_n$ of the dynamical variables that is independent of the time. From this fact he shows that the coefficients $x_{mn}$'s in the expansion of $x \psi_n$

$$
x \psi_n = \sum_m x_{mn} \psi_n, \quad (34)
$$

where $x$ is a function of dynamical variables that does not involve time explicitly, are of the form

$$
x_{mn} = \xi_{mne}^{i(W_m - W_n) t / \hbar}, \quad (35)
$$

where $\xi_{mn}$'s are constants.

$^{16)}$ See Sec. 3, §2 of the present chapter for the relativistic generalization of the wave equation.
Dirac says that the form of $x_{mn}$ given in Eq. (35) is "as in Heisenberg's theory". That is to say, by assuming Eq. (30) in consideration of the fact that time and energy make a pair of canonical variables in the Hamilton equation of motion, Dirac fills up, with respect to the time dependence, the mathematical equivalence between matrix mechanics and wave mechanics shown by Schrödinger in "the paper on the relationship".

§6. Dirac's consideration of Einstein's theory of transition probability

Dirac treats next transition probabilities of absorption and induced emission, by applying the wave equation obtained by means of Eq. (30) to "an atomic system subjected to a perturbation from outside (e.g., an incident electromagnetic field)". This is the same aim as that of Schrödinger's fourth paper which we have described in Sec. 4, §4 of the last chapter. Schrödinger did not considered the initial condition, however, because he treated the perturbational effect of an external periodic field as a forced oscillation, so to speak, as we have noted in Footnote 61) at the end of the last chapter. On the other hand, Dirac solves the time-dependent wave equation by taking account of the initial condition.

If the perturbation is represented by a function $A$ of $p_r$'s, $q_r$'s and $t$, "the wave equation for the disturbed system" will be of the form

$$ (H - W + A)\psi = 0 \quad . \tag{36} $$

"The wave equation for the undisturbed system" is then

$$ (H - W)\psi = 0 \tag{37} $$

$H$ is assumed to be independent of $t$. The general solution of Eq. (37) is given by the "set of independent solutions" or the "set of eigenfunctions" in the form

$$ \psi = \sum_n c_n \psi_n \quad , \tag{38} $$

where $c_n$'s are constants.

Dirac tries to solve Eq. (36) by putting

$$ \psi = \sum_n a_n \psi_n \quad , \tag{39} $$
where the $a_n$'s are functions of $t$ only. He expands $A\psi_n$ in the form

$$A\psi_n = \sum_m A_{mn} \psi_m,$$

where the coefficients $A_{mn}$'s are functions of $t$ only and $A_{mn}^* = A_{nm}$, * denoting the imaginary conjugate. Since $H$ and $A$ commute with $a_n$ and there holds the relation $Wa_n - a_n W = i\hbar \dot{a}_n$ with respect to $W = i\hbar \partial / \partial t$, we find

$$i\hbar \dot{a}_m = \sum_n a_n A_{mn}. \quad (41)$$

Taking the imaginary conjugate of this equation, we get $N_m = a_m a_n^*$. Hence, putting $N_m = a_m a_n^*$, we have

$$i\hbar \sum_m \dot{N}_m = \sum_{nm} (a_n^* A_{mn} a_n - a_n A_{nm} a_m) = 0. \quad (43)$$

From this, Dirac assumes that $|a_m|^2$ is the number of atoms in the $m$th state", because Eq. (43) then "makes the total number of atoms remain constant".

If the perturbation consists of incident electromagnetic radiation moving in the direction of the $x$-axis and plane-polarized with its electric vector in the direction of the $y$-axis, the perturbing term $A$ in the Hamiltonian is given, neglecting relativistic affects, by $A = (K/c)\eta$, where $\kappa$ and $\eta$ are the component in the direction of the $y$-axis of the vector potential and the total polarization respectively. Expanding $\eta \psi_n$ and $\eta \dot{\psi}_n$ in the form of Eq. (34), we have $\eta_{mn} = i\omega_{mn} \eta_{mn}$ for their respective constants $\eta_{mn}$ and $\eta_{mn}$ corresponding to $\xi_{mn}$ in Eq. (35), where $\omega_{mn} = (W_m - W_n) / \hbar$. We then get from Eq. (41)

$$i\hbar \dot{a}_m = \sum_n a_n \kappa \eta_{mn} e^{i\omega_{mn} t}. \quad (44)$$

Dirac integrates this equation by iterative approximations in $\kappa$ on the initial condition that $a_m = c_m$ at $t = 0$, and gives $a_m$ at $t = T$ in the form $a_m = c_m + c'_m + c''_m$, where $c'_m$ and $c''_m$ denote terms of first and second orders in $\kappa$ respectively.

"To obtain effects that are independent of the initial phases of atoms", Dirac substitutes $c_m e^{i\gamma_m}$ for $c_m$ and averages over all values of $\gamma_m$. When this averaging procedure is applied to $N_m = a_m a_n^*$, the first-order terms of $N_m$, namely $c'_m c'_n$ and $c_m c_m^*$, vanish. It is found consequently that the change
\( \Delta N_m \) in \( N_m \) during the time from \( t = 0 \) to \( t = T \) is equal to

\[
\Delta N_m = \left( \frac{1}{\hbar c} \right)^2 \sum_m \left\{ |c_n|^2 - |c_m|^2 \right\} \left| \eta_{nm} \int_0^T \kappa(t) e^{i\omega_{mn} t} \, dt \right|^2 . \quad (45)
\]

\( \Delta N_m \) gives "the increase in the number of atoms in the state \( m \)". Each term in this expression of \( \Delta N_m \) "may be regarded as due to transitions between the state \( m \) and the state \( n \)".

If the radiation from \( t = 0 \) to \( t = T \) is resolved into its harmonic components, the intensity of frequency \( \nu \) per unit frequency range is

\[
I_{\nu} = \left( \frac{2\pi \nu^2}{c} \right)^2 \left| \int_0^T \kappa(t) e^{2\pi i\nu t} \, dt \right|^2 . \quad (46)
\]

If one averages over all the directions and states of polarization of the incident radiation and puts \( 2\pi \nu = \omega_{mn} \), the term \( \Delta N_m \) due to the transition between the states \( m \) and \( n \) becomes

\[
\begin{aligned}
(2\pi/3\hbar^2 c) \{ |c_n|^2 - |c_m|^2 \} |P_{nm}|^2 I_{\nu} ; \\
|P_{nm}|^2 = |\xi_{nm}|^2 + |\eta_{nm}|^2 + |\zeta_{nm}|^2 ,
\end{aligned}
\]

being \( \xi \) and \( \eta \) the components of the polarization in the directions of the \( x \)- and \( z \)-axes respectively. "Thus one can say that the radiation has caused \( (2\pi/3\hbar^2 c) |c_n|^2 |P_{nm}|^2 I_{\nu} \) transitions from state \( n \) to state \( m \) and \( (2\pi/3\hbar^2 c) |c_m|^2 |P_{nm}|^2 I_{\nu} \) transitions from state \( m \) to state \( n \), the probability coefficient for either process being

\[
B_{n \rightarrow m} = B_{m \rightarrow n} = \left( \frac{2\pi}{3\hbar^2 c} \right) |P_{nm}|^2 , \quad (48)
\]

in agreement with the Einstein theory".

In this way, Dirac succeeds in "accounting for the absorption and stimulated emission",\(^{17}\) and in "showing that the elements of the matrices representing the total polarization determine the transition probabilities". However, his theory "cannot take spontaneous emission into account". Dirac considers that for this purpose it is necessary to have "a more elaborate theory involving the positions of the various atoms and interference of their individual emissions". He discusses that "the effects will depend upon whether the atoms

\(^{17}\)The "stimulated emission" means the induced emission in the present-day terminology.
are distributed at random, or arranged in a crystal lattice, or all confined in a volume small compared with a wavelength”, and says that “the last alternative mentioned, which is of no practical interest, appears to be the simplest theoretically”.

In solving the time-dependent wave equation, that is, the time dependent Schrödinger equation in the present-day nomenclature, Dirac sets the initial condition $a_m = c_m$, and takes average over all the possible phases of $c_m$ in consideration of an aggregate of atoms, because he supposes that $\sum_m N_m$ represents the total number of atoms. This differs from the procedure seen in present-day textbooks on quantum mechanics, in which one considers a single atom and sets the initial condition $a_m = \delta_{mk}$, that is, $\psi = \psi_k$. On the average over the initial phases, Dirac makes the following discussions.

If initially all the atoms are in the ground states, Eq. (45) holds without the averaging process. If, on the other hand, some atoms are initially in an excited state, we may suppose that they were brought into this state by radiation incident on the atoms before the time $t = 0$. The effect of the subsequent incident radiation must then depend on its relationships with the earlier incident radiation. If we do not wish the earlier radiation to appear explicitly in the calculation, we must suppose that it impresses certain phases on the atoms it excites. Thus, Dirac concludes that “the initial phases are of real physical importance, and in consequence the Einstein coefficients are inadequate to describe the phenomena except in special cases”, and states that “it should be observed that we get the simple Einstein results only because we have averaged over all initial phases of the atoms”.

In other words, Dirac interpreted the Einstein theory of transition probabilities as a kind of statistical theory of an aggregate of atoms. In this respect, Dirac’s viewpoint shown in the paper under consideration was like that presented by Bohr-Kramers-Slater. Though any mention of it was not yet made in Dirac’s paper under consideration, Born gave as we have seen in Sec. 5 of the last chapter, the corpuscular-theoretical interpretation of the wave function as the probability amplitude, to derive the formula for the probability of scattering of particles with the use of the time-independent Schrödinger equation, in his provisional paper and full paper received two months and one month before Dirac’s paper mentioned, respectively.

As will be described in §1 of the following section, after the presentation of the time-dependent wave equation by Schrödinger, Born treated the time

---

18) Cf. Sec. 3, §§1 and 2, Chap. 5, Vol. II.
variation of the wave function to give the probability, that is, the wave function as probability amplitude, and derived the transition probability. This solved the difficulties discusses by Dirac over the transition probabilities of absorption and induced emission. The problem of the transition probability of spontaneous emission came to be solved when quantum mechanics was applied after its establishment to the radiation field. This process will be described in Sec. 3, §3 of the present chapter.

2. The Unification as Quantum Mechanics

§1. Time variation of $\psi$ and transition probabilities

In the paper "The adiabatic Principle in Quantum Mechanics", received on Oct. 16, 1926, Born attempts to "substantiate by new applications the usefulness" of the standpoint presented in his previous papers. From this standpoint, the coefficient $c_n$ of the eigenfunction $\psi_n$ in the expansion of a wave function $\psi$ in terms of the set of eigenfunctions is interpreted as giving the state probability $|c_n|^2$. Namely, $|\psi|^2$ is grasped based on the idea, which later becomes one of the fundamental concepts in quantum mechanics, that it gives the probability of finding a particle. Now, Born says that "this standpoint consists in seeing in quantum mechanics the fusion of mechanics and statistics in the meaning that the new mechanics does not, as the old does, answer the question 'how a particle moves', but the question 'how it is probable that a particle moves in a given way'".

Born mentions that in the fourth paper "also Schrödinger deals with the squared amplitude of the wave and introduces for it the expression 'weight function', which comes already very close to the terminology of the statistics". Born points out as follows, too — "Mr. Schrödinger has adopted a standpoint which deviates from his earlier one and, independently of my arguments, come to an interpretation which is similar to mine". However, Born has "the impression that Schrödinger strives to describe quantum processes by a causal continuum theory in the classical sense". To Born "this seems certainly

20) M. Born, Zeit. f. Phys. 37 (1926), 863; 38 (1926), 803 [the same as Refs. 52) and 53) in Chap. 2 of the present volume].
21) Cf. §6 of the present section.
22) E. Schrödinger, Ann. d. Phy. 81 (1926), 102 [same as Ref. 47) in Sec. 5, §1, Chap. 2 of the present volume].
impossible”. Born says that “it is of course without doubt that Schrödinger’s formulae are right”, and “the mathematical getting-through of the quantum laws, which he has done, is certainly the key to their true understanding”. He considers that “fortunately, the formulae, which indeed refer up to now almost exclusively to stationary states, by no means lead to Schrödinger’s interpretation”.

Born continues to argue that “it seems rather indispensable here to bring in the basic ideas of the matrix form of the theory founded by Heisenberg”, and that “these ideas have grown up from the natural description of atomic processes by ‘quantum jumps’, and stress the classically and geometrically incomprehensible character of these processes”. These words of Born show that he notices the situation, mentioned in the last chapter, that matrix mechanics developed on the premise of the expression for transition probabilities, which it could introduce because of its phenomenological character.

Born says then that “it is firm that for stationary states both the forms of the theory arrive at the same result”, and “it is now in question how nonstationary processes are to be treated”. He considers that “about this, Schrödinger’s formalism proves to be essentially more handy, provided that one interprets it in Heisenberg’s sense”. Thus, Born “wants to support a fusing of the two ways of looking at things, by which each comes to quite a determined function”.

He takes now the case of an atom with plural electrons. Its nucleus is assumed to be infinitely heavy. Denoting the Cartesian coordinates of all the electrons simply by \( x \), Born writes “Schrödinger’s differential equation” as

\[
\Delta \psi - \left( \frac{8\pi^2 \mu}{\hbar^2} \right) U(x, t) \psi - \left( \frac{4\pi^2 \mu}{\hbar} \right) \frac{\partial \psi}{\partial t} = 0, \tag{49}
\]

where \( \mu \) is the electron mass and \( U(x, t) \) is the potential energy which depends in general on the time \( t \). If \( U(x, t) \) does not depend on the time and is a function \( U(x) \) only of \( x \), “there exist stationary states, whose solutions correspond to the form

\[
\psi(x, t) = \psi_m(x) e^{2\pi i W_m t / \hbar}. \tag{50}
\]

For simplicity it is assumed that the system is not degenerate. The \( \psi_m \)'s are orthogonal and normalized to unity, so that

\[
\int \psi_n(x) \psi_m^*(x) dx = \delta_{n,m}. \tag{51}
\]

\(^{23}\)\( \Delta \) used here by Born denotes the Laplacian operator in the configuration space.
By "superposition" of the functions given in Eq. (50), a "general solution" of Eq. (49) is obtained in the form

$$\psi(x, t) = \sum_{m} c_m \psi_m(x) e^{2\pi i W_m t / \hbar}. \tag{52}$$

Born asks then the question "what this function physically means". It also concerns the question "in which cases the results of wave mechanics can be interpreted in the words of quantum jumps". He says that "we are certain that numerous phenomena can simply be interpreted with the help of the concepts of particles, and we attempt to bring this into harmony with wave mechanics". He thus "holds on to Bohr's picture that an atomic system is always only in a stationary state". As we have seen in and after Chap. 2, Vol. II, the proposal in the Bohr model of the existence of such stationary states and of transitions between stationary states provided the basis for the development of the old quantum mechanics.

Born argues that "we can even assume that we know occasionally, by observation at a moment with certainty, that the atom is in the mth state", and further that "in general, however, we shall at a moment only know that, on the basis of the history and the existing physical conditions, a certain probability exists that the atom is in the mth state". Born "asserts then that as a measure this state probability is given by the quantity

$$|c_m|^2 = \left| \int \psi(x, t) \psi_m^*(x) dx \right|^2. \tag{53}$$

Thus, the important relation that constitutes the ground of the present-day quantum mechanics is here presented.

Born then considers the case in which the potential energy $U(x, t)$ consists of the inner potential energy $U(x)$ of the atom and an external perturbing one $F(x, t)$, so that he puts

$$U(x, t) = U(x) + \kappa F(x, t), \tag{54}$$

where $\kappa$ denotes the parameter of perturbation. He supposes that the perturbation acts during the time from $t = 0$ to $t = T$ and that

$$\psi(x, 0) = \psi_n(x). \tag{55}$$

---

24) Born says in a footnote that "in a gas left at a sufficiently low temperature all atoms with few exceptions are in the normal states".
This means the initial condition that $c_m = \delta_{nm}$ at $t = 0$. Born solves the differential equation (49) on this initial condition by means of the iterative method in $\kappa$, and gets

$$\psi_n(x, t) = \sum_m b_{nm} \psi_m(x) e^{2\pi i k W_m t / h}, \quad (t > T), \quad (56)$$

where

$$\begin{align*}
\begin{cases}
    b_{nm} = f_{nm}(T) \\
    f_{nm}(t) = \sum_{p=0}^{\infty} f_{nm}^{(p)}(t) \kappa^p, \quad f_{nm}^{(0)}(t) = \delta_{nm}, \\
    f_{nm}^{(p)}(t) = \left( \frac{2\pi i}{\hbar} \right) \int_0^t \sum_k f_{nk}^{(p-1)}(t') F_{km}(t') e^{2\pi i \nu_k t'} dt', \\
    F_{nm}(t) = \int F(x, t) \psi_n \psi_m^*(x) dx, \quad h\nu_{km} = W_k - W_m,
\end{cases}
\end{align*} \quad (57)$$

$F_{nm}(t)$ given above is "the matrix relevant to $F(x, t)$".

From the "basic hypothesis", $|b_{nm}|^2$ is the probability that the system is found in the state $m$ at $t > T$, so that "one can call $|b_{nm}|^2$ the transition probability from $n$ to $m". If the perturbation is due to an external electric field $E(t)$, one has $\kappa F_{nm}(t) = -M_{nm} E(t)$ in terms of "the matrix relevant to $M(x)$", the component of the electric dipole moment parallel to the electric field. Putting

$$f(\nu) = \int_0^T E(t) e^{-2\pi i \nu t} dt, \quad (58)$$

one gets in the first approximation

$$|b_{nm}|^2 = \left( \frac{2\pi}{\hbar} \right)^2 |M_{nm}|^2 |f(\nu_{nm})|^2 \quad (n \neq m). \quad (59)$$

In order to find "Einstein's transition probability per unit time" from this result, Born considers "natural light to be an irregular sequence of many wave pulses", and takes as $T$ the duration of each individual pulse. Denoting the average over these pulses by $\langle \rangle$ and the number of pulses per unit time by $N$, he shows that

$$\langle E^2(t) \rangle = \left\langle \frac{1}{T} \int_0^T E^2(t) dt \right\rangle = N \int_{-\infty}^{\infty} \langle |f(\nu)|^2 \rangle \nu$$
by making use of the relation $\int_0^T E^2(t) dt = \int_0^\infty \langle |f(\nu)|^2 \rangle d\nu$. As the radiation density $u$ is given by $u = \langle E^2(t) \rangle / 4\pi = \int_0^\infty \rho(\nu) d\nu$, $\rho(\nu)$ turns out to be $\rho(\nu) = (N/2\pi) \langle |f(\nu)|^2 \rangle$. Denoting the electric dipole moment of the atom by $M$, one has the relation $E^2 M^2 = E^2(M^2/3)$ after averaging over the directions of the atom. $B_{nm} \rho(\nu_{nm})$ is equal to the sum of Eq. (59) over the $N$ pulses, and is therefore given by

$$B_{nm} \rho(\nu_{nm}) = \left( \frac{4\pi^3}{3\hbar^2} \right) |M_{nm}|^2 \rho(\nu_{nm}).$$

"It is valid naturally that $B_{nm} = B_{mn}$".

In this way, Born succeeded in showing that the transition probability introduced by Einstein in 1916 for "in-radiation" (absorption and induced emission of radiation) is given in the same form as was assumed in matrix mechanics, and advanced a decisive step to the "fusing" of matrix mechanics and wave mechanics. The framework of quantum mechanics, as the essentialistic theory unifying matrix mechanics and wave mechanics, was thereby provided.

The framework of this unified theory is based on the equality of the two mechanics with regard to periodic phenomena, and is also based on the fact that the phenomenological as well as corpuscular-theoretical matrix mechanics, on one hand, is able to assume transition probabilities, but unable to treat aperiodic phenomena such as collision processes, while the substantialistic wave mechanics, on the other hand, is able to treat collision processes and leads to results describing well the scattering experiments, if the wave function is related to the probability of finding a particle, by the introduction of the corpuscular-theoretical viewpoint, as was shown by Born in his foregoing paper. The unification under consideration could be achieved only if the wave function obeys a first order differential equation with respect to time, as we have noted in Sec. 5, §4 of the last chapter, because in that case the initial value of the probability amplitude alone determines its value at any later time. By making use of the time-dependent wave equation found by Schrödinger, Born advanced a step to the unification.

As we have seen in the last chapter, Schrödinger showed in his "paper on the relationship" that matrix elements are determined by wave functions and vice versa, but for the time dependence of matrix elements applied the same assumption as that used in matrix mechanics. The time-dependent wave equation presented by Schrödinger was what resolved this problem. In fact, Dirac presented independently the time-dependent wave equation, in his paper
received two months later than Schrödinger’s fourth paper, and derived the
time dependence of matrix elements, as we have described in §5 of the last
section. Though matrix mechanics and wave mechanics were thus unified, it
was still not clear how the wave function itself should be put into the formalism
of matrix mechanics. We shall take up this problem in §3 of the present section.

In Born’s paper under consideration, also, the transition probability for
Einstein’s “out-radiation”, that is, spontaneous emission of radiation, was not
treated, just like Dirac’s paper mentioned in §6 of the last section. By the
way, Born discussed in this paper “the reversibility of transition probabilities”,
showing that the same type of differential equation as Eq. (49) is obtained for
the complex conjugate $\psi^*$ of $\psi$, if the transformation $t \to -t$ is applied to
the complex conjugate of Eq. (49). He showed in addition that “in the case
of infinitely slow perturbation there is no probability of a quantum jump”, by
considering perturbations such that $F(x,t) = 0$ for $t \geq T$ as well as $t \leq 0$ and
taking the limit $T \to \infty$. According to Born “this law corresponds well with
the classical law of adiabatic invariance of action integral”.

§2. Wave function as a “state quantity” in Hilbert space

London showed, in his paper “Angle Variable and Canonical Transformation
in Wave Mechanics”\(^\text{25}\) received on Sep. 19, 1929, that “the application of the
transformation theory in matrix mechanics to Schrödinger’s eigenvalue theory
leads to a very generalized view of wave mechanics”.

In this paper, London notices the fact that in wave mechanics “the q-space
is exclusively used”, while in classical mechanics the solution of a mechanical
problem is obtained by means of a transformation of the canonical variables.
The method of canonical transformation is also used in matrix mechanics, and
a transformation from canonical coordinates $p, q$ as matrices to new canonical
coordinates $P, Q$ is called a canonical transformation, if it keeps the condition

$$pq - qp = \begin{pmatrix} \hbar \\ i \end{pmatrix} \mathbf{1}$$

(61)
invariable, where $\mathbf{1}$ denotes the unit matrix.\(^\text{26}\)


\(^{26}\)Following Dirac, London uses $\hbar$ for $\hbar/2\pi$, which we denote by $\hbar$ to avoid confusion.
London writes "the differential equation of wave mechanics" as
\[
[H \left( q, -i\hbar \frac{\partial}{\partial q} \right), \Psi(q)] = E\Psi(q),
\]
according to the notation in Schrödinger's "paper on the relationship.\(^{27}\) In wave mechanics, solutions of Eq. (62) are sought for by means of transformations of the kind \( Q = f(q) \) in the \( q \)-space. London points out that "these correspond to 'point transformations' in the mechanics". He intends "to generalize point transformations in wave mechanics by analogy to canonical transformations", and studies transformations of the form
\[
\begin{aligned}
Q &= F \left( q, \frac{\partial}{\partial q} \right), \\
\frac{\partial}{\partial Q} &= G \left( q, \frac{\partial}{\partial q} \right),
\end{aligned}
\]
which transform \( q \) and \( \partial/\partial q \) simultaneously. \( F \) and \( G \) cannot be arbitrary functions independent of each other, and "for a canonical transformation" the differential operator determined by \( G \) should represent differentiation with respect to \( Q \) introduced by \( F \).

"In accordance with the invariance of the commutation rule", London demands the invariance of the operator
\[
\left( \frac{\partial}{\partial q} \right) q - q \left( \frac{\partial}{\partial q} \right) = 1
\]
for the transformation given by Eq. (63). For such canonical transformations in wave mechanics, he takes "the representations in matrix mechanics (not in classical mechanics!), in which \( p \) is replaced everywhere by \( (\hbar/i)\partial/\partial q \)."

As we have seen in Sec. 2, §4, Chap. 1 of the present volume, Born-Heisenberg-Jordan showed that the transformation \( Q = SqS^\dagger, P = SpS^\dagger \) of \( q, p \) by a unitary matrix \( S \) keeps the commutation relation (61) invariant, and named this transformation a canonical transformation in matrix mechanics (\( \dagger \) means the unitary conjugate). On the other hand, in classical mechanics Jacobi's expression \( Q = (\partial/\partial P)S(q, P), p = (\partial/\partial q)S(q, P) \) with the generating function \( S(q, P) \) is known as a canonical transformation. London shows, in another paper,\(^{28}\) that also in matrix mechanics "correspondingly with Jacobi's

\(^{27}\)Cf. Eq. (77) in Sec. 3 §2, Chap. 2 of the present volume.

expression”, the relation $Q = (\partial/\partial P)S(q, P), \ p = (\partial/\partial q)S(q, P)$ with a real matrix function $S(q, P)$ of $q$ and $P$ makes a canonical transformation. Making use of this result, London puts as a canonical transformation in wave mechanics

$$Q = \frac{\partial}{\partial P}S\left(q, \frac{\hbar}{i} \frac{\partial}{\partial Q}\right), \quad \frac{\hbar}{i} \frac{\partial}{\partial q} = \frac{\partial}{\partial q}S\left(q, \frac{\hbar}{i} \frac{\partial}{\partial Q}\right)$$

(65)

according to the method mentioned above. $S$ in Eq. (65) is supposed to be adequately symmetrized, similar to $S(q, P)$ which should be adequately symmetrized with respect to $q$ and $P$.

Now, all the eigenvalue problems which are derived from Eq. (62) by canonical transformations of the type given in Eq. (65) have common eigenvalues. London says that to show this it is “easier” to use another representation of canonical transformation given by Jordan,\(^9\) which is equivalent to Eq. (65). This representation is derived on the supposition that $S(q, P)$ is given by a sum of products of a function of $q$ only with one of $P$ only, as in $S(q, P) = \sum_n f_n(q)q_n(P)$. It is given by

$$q = T^{-1}Q, \quad \frac{\partial}{\partial q} = T^{-1}\left(\frac{\partial}{\partial Q}\right) T,$$

(66)

in terms of an arbitrary real differential operator $T(Q, \partial/\partial Q)$ and its reciprocal operator $T^{-1}$. Letting Eq. (62) to be the original eigenvalue problem, London writes the canonically transformed eigenvalue problem as

$$H^* \left[\left(Q, -i\hbar \frac{\partial}{\partial Q}\right), \Psi^*(Q)\right] = E^* \Psi^*(Q),$$

(67)

with the mark * of discrimination. $H^*$ is given by

$$H^* \left(Q, -i\hbar \frac{\partial}{\partial Q}\right) = T^{-1}H \left(Q, -i\hbar \frac{\partial}{\partial Q}\right) T.$$  

(68)

Consequently, by operating $T$ on Eq. (67) from the left-hand side, one gets

$$\left[H \left(Q, -i\hbar \frac{\partial}{\partial Q}\right) T, \Psi^*(Q)\right] = E^*[T, \Psi^*(Q)].$$  

(69)

If one puts

$$\left[T \left(Q, \frac{\partial}{\partial Q}\right) T, \Psi^*(Q)\right] = \Psi(Q),$$  

(70)

the differential equation (69) is identical with the original differential equation, so that it has the same sequence of eigenvalues as that of the original equation.

London interprets the "meaning" of this result as follows. One has in a domain a mapping $H$ which transfers every object $x$ to another object $y$ in the domain. Besides, one has another mapping $T$ which maps the whole domain together with its mapping $H$ to a new domain, so that $x$ goes over to $x^*$ and $y$ to $y^*$. "The mapped mapping $H^*$ transfers therefore $x^*$ to $y^*$". By making use of the detour $x^* \rightarrow x \rightarrow y \rightarrow y^*$ instead of $x^* \rightarrow y^*$ one gets the well-known representation of the transformation of a transformation

$$H^* = T^{-1}HT.$$ (71)

From Eq. (69) it is seen that "the object $x$ is Schrödinger's new state quantity $\Psi$". "Oscillating phenomena of $\Psi$ are described by a series of eigenfunctions $\Psi_k^*$", and "the operator $T$ transfers this series term by term to another system of eigenfunctions $\Psi_j^n$". Namely, if one puts

$$\Psi_k(Q) = \sum_j T_{jk} \Psi_j^*(Q),$$ (72)

one has the relation

$$T_{jk} = \int \overline{\Psi_j^* \Psi_k} dQ = \int \overline{\Psi_j^*} \left[T \left(Q, \frac{\partial}{\partial Q}\right), \Psi_k^*\right] dQ,$$ (73)

where $\overline{\Psi_j^*}$ denotes the complex conjugate of $\Psi_j^*$, and so on. "One therefore gets as the development coefficient $T_{jk}$ just the one which Schrödinger calls the matrix element $T_{jk}$".

Making use of Eq. (71) and the relations $\int \overline{\Psi_k} \Psi_{k'} dQ = \delta_{kk'}$ and $\int \overline{\Psi_j} \Psi_{j'} dQ = \delta_{jj'}$, one gets for the matrix $T$, whose elements are $T_{jk}$, and its "transposed matrix" $\overline{T}_{30}$

$$\overline{TT} = 1.$$ (74)

London concludes that Eq. (73) "characterizes the mapping given by Eq. (72) as a 'rotation' (in Hermite's meaning) of the 'system of coordinate axes' spanned by the orthogonal eigenfunctions in the Hermite space of infinitely many dimensions". 31)
In this way, London showed that the wave function is a "state quantity" in the Hilbert function space spanned by the eigenfunctions as the "axes". Thus it was pointed out for the first time that in quantum mechanics a state of a physical system is represented by a vector in the Hilbert space. As we have seen in Sec. 2, §4, Chap. 1 of the present volume, Born-Heisenberg-Jordan introduced canonical transformations in matrix mechanics, and showed that the diagonalization of a matrix by a canonical transformation is equivalent to the transformation of a quadratic form into it in the principal axes.

In matrix mechanics, however, the principal axes as vectors were used for the convenience of calculation, and were not examined for their physical meaning, since the main interest was in matrix elements, which were introduced phenomenologically as quantities relating transitions between stationary states. In wave mechanics, which introduced substantialistically wave functions as describing "oscillating phenomena" of an atom as in London's saying, the main interest were in stationary states themselves. London opened the way to interpreting physically the principal vectors in matrix mechanics by applying of canonical transformation in matrix mechanics to wave mechanics.

§3. The concept of probability for matrix elements of transformation

In matrix mechanics, it was supposed that, when the Hamiltonian put in the matrix form is diagonalized by means of a canonical transformation, its diagonal elements give the energies of the discrete stationary states in the periodic motions of the system. In wave mechanics, on the other hand, it was supposed that, when a wave function, as a continuous function of the position coordinates, is an eigenfunction of Schrödinger's differential equation, the eigenvalue gives the energy of a stationary state of the system. In addition, the wave function is interpreted as being connected with the probability of finding the system in the stationary state. This interpretation then makes it possible for wave mechanics to treat aperiodic motions such as that in collision processes. However, there still remained the problem of how these relations should be seen from the standpoint of matrix mechanics.

Dirac studied this problem in his paper "The Physical Interpretation of Quantum Dynamics"\(^{32}\) received on Dec. 2, 1926, citing Schrödinger's fourth

paper\textsuperscript{33}) and Born's paper on collision problem\textsuperscript{34}) In his paper under consideration, Dirac says about Schrödinger's fourth paper that Schrödinger "has provided new ways of obtaining physical results from the theory, based on the assumption that the square of the amplitude of the wave function can in certain cases be interpreted as a probability". Namely, Dirac sees Schrödinger's supposition that the square of the wave function should be the "weight function" in the configuration space as introduction of the probability interpretation, just like Born who points out in his paper mentioned in §1 of the present section that "Schrödinger has come to an interpretation which is similar to" Born's.

Dirac finds the clue to tackle this problem in Heisenberg's of results of which he is informed "before publication". Heisenberg's results quoted here refer to those given in his paper "Fluctuation Phenomena and Quantum Mechanics"\textsuperscript{35}) which was received on Nov. 6, 1926, about one month earlier than Dirac's paper under consideration. Heisenberg says in the mentioned paper that "the great physical meaning of fluctuation phenomena is based on the fact that they are relevant to the discontinuities appearing with processes in small spaces and short times".

Heisenberg continues to say that "according to Einstein, for example, Brownian motions can be considered as the direct consequence of the atomic structure of matter, the fluctuations in energy and momentum of the cavity radiation lead directly lead to Einstein's idea of light quanta, and the energy fluctuation such as in a crystal lattice is in closest connection with the existence of discrete stationary states of a mechanical system". These are all related to matters which played important roles, as we have described in detail in Chap. 1, Vol. II of the present book, in the development of the quantum theory in the substantialistic stage.

Heisenberg considers then that "because the theory of mechanical systems has been made accessible to a quantitative treatment by quantum mechanics, a very close connection will have to exist between quantum mechanics and each of the above-mentioned kinds of fluctuation phenomena". He says that a closer investigation of such connections seems to him useful, "because recently from many sides doubts have been loud about the existence of the discontinuity". As

\textsuperscript{33})E. Schrödinger. Ann. d. Phys. \textbf{81} (1926), 109 [cf. Sec. 5, §1, Chap. 2 of the present volume].
\textsuperscript{34})M. Born, Zeit. f. Phys. \textbf{37} (1926), 863; \textbf{38} (1926), 803 [cf. Sec. 5, §3, Chap. 2 of the present volume].
we have described in Sec. 1, §1 of the present section, Heisenberg attached, in his paper about three months earlier on “Many-body Problem and Resonance in Quantum Mechanics”, importance to “one of the aspects of quantum mechanics that quantum mechanics is based on the corpuscular idea” from the viewpoint that the wave theory of light cannot explain completely all experiments in optics, though he appreciated the success of wave mechanics. “Quantum mechanics” in Heisenberg’s these words means matrix mechanics.

For the study of the connection of matrix mechanics with the fluctuation phenomena, Heisenberg treats “the mathematical part” from this paper on resonance. That is, as we have seen in Sec. 1, §2 of the present chapter, the energy of a system consisting of two identical atoms \(a\) and \(b\) is degenerate for the two cases, in one of which atom \(a\) is in state \(m\) and atom \(b\) in state \(n\), and in the other of which, conversely, atom \(a\) is in \(n\) and atom \(b\) in \(m\). The interaction between the two atoms resolves the degeneracy. The two-dimensional matrix \(S\) of the canonical transformation to diagonalize the interaction energy is given in the perturbation calculation as Eqs. (5a) and (5b) in the present chapter. If the energy matrix of atom \(a\) is denoted by \(W^a\), then in the perturbed system it is given by \(W'^a = S^{-1}W^aS\). Consequently, the diagonal and off-diagonal elements of \(W'^a\) are equal to \((E_n + E_m)/2\) and \((E_n - E_m)/2\), respectively, where \(E_n\) means the energy of the atom in state \(n\) and \(E_m\) that in state \(m\). When \(W'^a\) is averaged over time, the off-diagonal element becomes zero on account of its oscillation in time, the diagonal one alone remaining finite. One can arrive at the same result, also, by considering that the atoms \(a\) and \(b\) “resonate” with each other, and to them “energy jumps occur with a certain frequency discontinuously”, so that on the time average their energies are both equal to \((E_n + E_m)/2\) by reason of the symmetry between them. Heisenberg says that “in other words, the ‘light quantum’ \((E_n - E_m)\) time and time again goes over from atom \(a\) to atom \(b\), and “on the average the light quantum will be in atom \(a\) for a half of the time, and in atom \(b\) for the other half”.

Generally, also in the case of the matrix \(f(W^a)\) given in terms of a function \(f(x)\), the time average of the diagonal element alone of \(S^{-1}f(W^a)S\) remains different from zero, being equal to \(\{f(E_n) + f(E_n)\}/2\). Heisenberg asks the question “whether it is possible by collision processes with one of the two atoms, for example, to realize experimentally energy values, which are somewhere

\[36\) Cf. Ref. 2) in Sec. 1, §1 of the present chapter.
\[37\) Heisenberg says in a footnote that one should perhaps better use “sound quantum”, because it concerns states of mechanical oscillation.
between $E_n$ and $E_m$, and answers "no". The factor $1/2$ in the expression given just above comes from the squared value of the elements of the canonical transformation matrix $S$, and this factor $1/2$ is interpreted as the average frequency of atom $a$ to be in the state $n$ or $m$. Heisenberg generalizes this consideration, thus, to the case of a many-body system consisting of identical subsystems, as follows.

"If one denotes the states of some perturbed mechanical system with $\alpha'$, and the states of the unperturbed system with $\beta'$, then in the sense of the idea of discontinuity the canonical transformation means the following: If the perturbed system is in the state $\alpha$, $|S_{\alpha\beta}|^2$ is the probability that the system is caught in the state $\beta$ (in the case of collision processes, sudden stop of the perturbation, etc.)." Because $S_{\alpha\beta}$ is an element of the canonical transformation matrix, one of course has $\sum\beta |S_{\alpha\beta}|^2 = 1$. The time averaged value $f_\alpha(W^a)$ in the state $\alpha$ of a function $f(W^a)$ of the energy $W^a$ of a subsystem $a$ is given by $\sum_\beta |S_{\alpha\beta}|^2 f(W^a_\beta)$, where $W^a_\beta$ is the value of $W^a$ in the state $\beta$. In this way, Heisenberg attempts to connect the probability concept with the matrix elements of a canonical transformation.

It is interesting to note that Heisenberg's idea of "the probability that the system is caught in the state $\beta$ (in the case of collision processes, sudden stop of the perturbation, etc.)" is similar to Born's idea given in his papers on collision problems. As we have described in Sec. 5, §4, Chap. 2 of the present volume, Born supposed in these papers that the square of the absolute value of each coefficient in the expansion of the wave function in terms of the eigenfunctions for null potential energy should be the frequency of the respective scattering, on the assumption that the interaction causing scatterings would vanish at infinite distances. These papers of Born on collision problems are not cited in Heisenberg's paper under consideration.\textsuperscript{38}

Though in Heisenberg's idea $|S_{\alpha\beta}|^2$ is supposed to be the probability of the event mentioned above, every matrix element was originally introduced in matrix mechanics as a quantity concerned with two periodic stationary states. On account of this, Born said in the full paper on collision problems that "in grasping mathematically the asymptotic behaviour of the coupled particle", he had "not succeeded with the matrix form of quantum mechanics", and

\textsuperscript{38}Born's provisional paper on collision problems was published in Vol. 37 of Zeit. f. Phys., and his full paper on collision problems in Vol. 38 of the same journal. Heisenberg's paper on resonance, from which "the mathematical part" for his study in the paper under consideration was taken, was published also in the same Vol. 38.
analyzed collision problems on the basis of wave mechanics (Cf. Sec. 5, §2, Chap. 2 of the present volume).

To make good use of Heisenberg’s idea, it would be necessary to generalize the matrix elements to cover aperiodic states. There remains also the question of how Heisenberg’s idea would be applied to a system consisting of nonidentical subsystems. In Dirac’s paper cited at the beginning of the present subsection, these themes are studied, and Heisenberg’s result is shown to be “capable of wide extensions”. Namely, as will be seen in the next subsection, Dirac makes it clear, in connection with Born’s result, that Heisenberg’s result “can be applied to any dynamical system, not necessarily one composed of two parts in resonance with each other, and to any dynamical variable, not necessarily one that can take only quantized values”.

§4. Matrices with continuous row and column and the δ-function

In the paper cited above,39) Dirac points out that “in most atomic problems the electrons are given to be initially in definite orbits”, so that one takes as the canonical variables “the initial values of the action variables (or other first integrals that define the orbits)”, but “there are certain problems for which the electrons are not initially in definite orbits”, and “to treat such problems we would require a matrix representation of the dynamical variables whose rows and columns refer to other constants of integration of the system than the action valuables”. As an example of such problems, Dirac mentions “the problem of the interaction of a β-particle emitted by a radioactive atom with the orbital electrons of the atom”. That is to say the scattering problem of β-ray. Dirac says that in this example “the β-particle is initially in the nucleus”,40) and “it would probably be convenient to have the matrix rows and columns referring to the co-ordinates of the β-particle at the time of emission”.

Thus, Dirac proceeds to set up “a theory of the more general schemes of matrix representation, in which the rows and columns refer to any set of constants of integration that commute”, including “a matrix representation with matrices that have continuous ranges of rows and columns instead of the usual discrete sets”. He says that “this theory may be regarded as a

39) P. A. M. Dirac, Ref. 32) of the present chapter.
40) At that time the neutron was not yet discovered, and the nucleus was thought to consist of protons and electrons.
development of Lanczos’s field theory”. “Lanczos’s field theory” refers to the theory that we mentioned in Sec. 4, §1, Chap. 1 of the present volume. Lanczos had in mind the correspondence of the commutation relation $pq - qp = (\hbar / i)1$ to the relation for the kernels of integral equations of the independent variables $s$ and $\sigma$, $(pq - qp)(s, \sigma) = \hbar E(s, \sigma)$. In this relation $E(s, \sigma)$ is the unit kernel, and corresponds to the $\delta$-function which Dirac introduces in his paper under consideration.

Next, Dirac considers a dynamical system of $u$ degrees of freedom, which has the first integrals of the equation of motion $\alpha_1, \alpha_2, \ldots, \alpha_u$ (action variables or otherwise). Each row or column of matrices can be labeled by specified values $\alpha_r$'s of $\alpha_r (r = 1, 2, \ldots, u)$. Thus, the elements of the matrix representing any dynamical variable $g$ is expressed by $g(\alpha_1, \alpha_2, \ldots, \alpha_u; \alpha''_1, \alpha''_2, \ldots, \alpha''_u)$, or by $g(\alpha'\alpha'')$ for brevity. Dirac calls $g(\alpha'\alpha'')$ the “$\alpha$-scheme of matrix representation”. $\alpha'$ and $\alpha''$ may be discrete or continuous. Dirac writes all formulae as though they can take only continuous ranges of values. If discrete sets of values do occur, necessary changes have to be made. The law of the multiplication of matrices $g_1$ and $g_2$ then reads

$$g_1g_2(\alpha'\alpha'') = \int g_1(\alpha'\alpha''')d\alpha'''g_2(\alpha''\alpha'''),$$  \hspace{1cm} (75)

where $d\alpha'''$ means $d\alpha''_1d\alpha''_2\cdots d\alpha''_u$.

With regard to matrices with continuous ranges of rows and columns, the unit matrix $1(\alpha'\alpha'')$ must satisfy the relation

$$\int 1(\alpha'\alpha''')d\alpha'''y(\alpha''\alpha'') = y(\alpha'\alpha''),$$  \hspace{1cm} (76)

for any matrix $y$. Saying that “we need the function $\delta(x)$ to express the elements of the unit matrix”, Dirac introduces the $\delta$-function. “$\delta(x)$ is defined by”

$$\delta(x) = 0 \quad (x \neq 0)$$

$$\int_{-\infty}^{\infty} \delta(x)dx = 1.$$  \hspace{1cm} (77)

“Strictly, of course, $\delta(x)$ is not a proper function of $x$, but can be regarded only as a limit of a certain sequence of functions”.
\( \delta(x) \) is taken to be an even function, so that \( \delta(-x) = \delta(x) \). For any regular function \( f(x) \) one sees that
\[
\int_{-\infty}^{\infty} \delta(a - x)f(x - b)dx = f(a - b). \tag{78}
\]
The differential coefficient \( \delta'(x) \) of \( \delta(x) \) is an odd function, that is, one has \( \delta'(-x) = -\delta'(x) \). One gets by integration by parts the formula
\[
\int_{-\infty}^{\infty} \delta'(a - x)f(x - b)dx = f'(a - b). \tag{78'}
\]
Equations (78) and (78') are valid also if \( f(x - b) \) is put equal to \( \delta(x - b) \), in the sense that their products by a function of \( b \) should be integrated with respect to \( a \) or \( b \). In terms of \( \delta \)-functions of such properties the unit matrix is expressed as
\[
1(\alpha' \alpha''') = \delta(\alpha'_1 - \alpha'''')(\alpha'_2 - \alpha'''') \cdots \delta(\alpha'_u - \alpha''''), \tag{79}
\]
which is denoted for brevity as \( \delta(\alpha' - \alpha'''') \).

The canonical transformation of \( g \) by a matrix \( b, G = bgb^{-1} \), is written as
\[
G(\alpha' \alpha'') = \int\int b(\alpha' \alpha''')d\alpha'''g(\alpha''' \alpha(4))d\alpha(4)b^{-1}(\alpha(4) \alpha''). \tag{80}
\]
"When we make a transformation of this kind, we can at the same time make any permutation of the rows of the new matrices \( G \) and the same permutation of their columns, without interfering with any of the conditions" that they should satisfy in matrix mechanics. "There is thus no one-to-one correspondence between the rows and columns of the new matrices and those of the original matrices". Dirac "therefore modifies the notation" and writes Eq. (80) as
\[
G(\xi' \xi'') = G(\xi'_1, \xi'_2, \ldots, \xi'_u; \xi''_1, \xi''_2, \ldots, \xi''_u)
= \int\int b(\xi' \alpha')d\alpha'g(\alpha' \alpha'')d\alpha''b^{-1}(\alpha'' \xi''). \tag{81}
\]
Namely, he means to specify the rows and columns of \( b \) or \( b^{-1} \) separately. "The \( \xi' \)'s may, in fact, take quite different ranges of values from the \( \alpha' \)'s, or one could have the \( \xi' \)'s taking only discrete sets of values while the \( \alpha' \)'s can take continuous ranges of values, or vice versa".
Then, Dirac writes Eq. (81) "in the simplified form"

\[ g(\xi', \xi'') = \int \int (\xi'/\alpha') d\alpha' g(\alpha' \alpha'') d\alpha''(\alpha''/\xi''). \]  

(81')

In this formula, the difference between \( g \) and the transformed \( g \) is expressed by the difference in the respective schemes of representation, while \( b(\xi' \alpha') \) and \( b^{-1}(\alpha'' \xi'') \) are written respectively as \( (\xi'/\alpha') \) and \( (\alpha''/\xi'') \), and are called "the transformation functions". Putting

\[
\begin{align*}
g(\alpha' \xi'') &= \int g(\alpha' \alpha'') d\alpha''(\alpha''/\xi''), \\
g(\xi' \alpha'') &= (\xi'/\alpha') d\alpha' g(\alpha' \alpha''),
\end{align*}
\]

(82a, 82b)

one can simply write Eq. (81') as

\[ g(\xi', \xi'') = \int (\xi'/\alpha') d\alpha' g(\alpha' \xi'') = \int g(\xi' \alpha'') d\alpha''(\alpha''/\xi''), \]  

(81'')

where \( g(\alpha' \xi'') \) and \( g(\xi' \alpha'') \) correspond to \( gb^{-1} \) and \( bg \) respectively.

The successive transformations by the matrices \( b_1 \) and \( b_2 \) are equivalent to a transformation by the matrices \( b_2 b_1 \). One has therefore for the transformation functions \( (\xi'/\alpha') \), \( (\kappa'/\xi') \) and \( (\kappa'/\alpha') \) respectively of \( b_1 \), \( b_2 \) and \( b_2 b_1 \)

\[
\begin{align*}
(\kappa'/\alpha') &= \int (\kappa'/\xi') d\xi' (\xi'/\alpha'), \\
(\alpha'/\kappa') &= \int (\alpha'/\xi') d\xi' (\xi'/\kappa').
\end{align*}
\]

(83a, 83b)

Now, among matrices in the \( \xi \)-scheme of representation there can be diagonal matrices, say \( \xi_r \)'s, such that the \( \xi_r \) has the matrix elements

\[ \xi_r(\xi', \xi'') = \xi_r \delta(\xi_1' - \xi'_1) \delta(\xi_2' - \xi'_2) \cdots \delta(\xi_u' - \xi'u'_u) = \xi_r \delta(\xi_1' - \xi'_1). \]  

(84)

The matrices \( \eta_r \)'s canonically conjugate respectively to the \( \xi_r \)'s should satisfy the relations that correspond to those in the case of matrices with discrete sets of values of \( \xi_r \)'s,

\[ \xi_r \eta_r - \eta_r \xi_r = \hbar, \quad \xi_r \eta_s - \eta_s \xi_r = 0, (r \neq s); \quad \eta_r \eta_s - \eta_s \eta_r = 0. \]
It can be shown that such a matrix $\eta_r$ has the matrix elements

$$\eta_r(\xi' \xi'') = -i\hbar \delta(\xi'_1 - \xi''_1) \cdots \delta(\xi'_{r-1} - \xi''_{r-1})$$

$$\times \delta'(\xi'_r - \xi''_r) \delta(\xi'_{r+1} - \xi''_{r+1}) \cdots \delta(\xi'_u - \xi''_u).$$ (84')

For the proof of this, it suffices to calculate from the $\xi_r$ and $\eta_r$ given above the $(\xi' \xi'')$ element of $(\xi_r \eta_r - \eta_r \xi_r)$ by integrations by parts using the properties of the $\delta$-function. As a result, one has for a single degree of freedom the formula$^{41}$

$$(\xi \eta - \eta \xi)(\xi' \xi'') = i\hbar \delta(\xi' - \xi'').$$ (85)

In this way, taking into consideration matrices that have continuous ranges of rows and columns, Dirac introduces the $\delta$-function to represent the unit matrix, and at the same time, in consideration of canonical transformations between matrices in different schemes of representation, adopts matrices with rows and columns separately labeled, making use of such particular notations as $(\xi'/\alpha')$. We summarize here Dirac's notations as follows.

$\alpha, \xi$ etc.: Dynamical variables, collectively representing $\alpha_r, \xi_r$ etc. ($r = 1, 2, 3, \ldots, u$). $\alpha', \xi'$, etc. are their specified values.

$g(\alpha' \alpha'') = g(\xi, \eta)(\alpha' \alpha'')$: The $(\alpha' \alpha'')$ element of a dynamical variable $g$ as a functions of the canonical variables $\xi$ and $\eta$ — the $\alpha$-representation.

$(\xi'/\alpha')$: The element of the matrix $b$ of the transformation from the $\alpha$-representation to the $\xi$-representation ($g \to bgb^{-1}$) — the transformation function.

$g(\xi' \alpha'')$: The $(\xi' \alpha'')$ elements of the matrix $bg$.

$g(\alpha' \xi'')$: The $(\alpha' \xi'')$ element of the matrix $gb^{-1}$.

§5. Wave function as a transformation function

Dirac calculates then the $\xi_r$ and $\eta_r$ given respectively by Eqs. (84) and (84'). The latter is calculated with the use of Eqs. (78') and (82) to be

$$\eta_r(\xi' \alpha') = \int \eta_r(\xi' \xi'')d\xi''(\xi''/\alpha') = -i\hbar \left( \frac{\partial}{\partial \xi'_r} \right)(\xi'/\alpha'),$$ (86a)

$^{41}$This formula is what corresponds to Lanczos’s equation $(pq-qp)(s, \sigma) = \hbar E(s, \sigma)$ shown in Sec. 4, §1, Chap. 1 of the present volume.
The Establishment of Quantum Mechanics

while the former is obtained using Eqs. (78) and (82) to be

$$\xi_r(\xi'\alpha') = \int \xi_r(\xi'\xi'')d\xi''(\xi''/\alpha') = \xi_r'(\xi'/\alpha').$$  \hspace{1cm} (86b)

He shows more generally that for a function $f(\xi_r)$ only of $\xi_r$ there holds the formula

$$f(\xi_r)(\xi'\alpha') = f(\xi_r')(\xi'/\alpha').$$  \hspace{1cm} (87)

Also, for any function $f(\xi_r, \eta_r)$ of $\xi_r$ and $\eta_r$ that is rational and integral in $\eta_r$, Dirac gives the theorem\(^\text{42)}\)

$$\begin{cases} 
  f(\xi_r, \eta_r)(\xi'\alpha') = f \left( \xi_r', -i\hbar \frac{\partial}{\partial \xi_r'} \right) (\xi'/\alpha'), \\
  f(\xi_r, \eta_r)(\alpha'\xi') = f \left( \xi_r', i\hbar \frac{\partial}{\partial \xi_r'} \right) (\alpha'/\xi'). 
\end{cases}$$ \hspace{1cm} (88a,b)

If a specified function $F(\xi_r, \eta_r)$ of this kind is diagonal in the $\alpha$-scheme of representation, one has from Eqs. (88a) and (82), by putting $F(\alpha'\alpha'') = F(\alpha')\delta(\alpha' - \alpha'')$,

$$F \left( \xi_r', -i\hbar \frac{\partial}{\partial \xi_r'} \right) (\xi'/\alpha') = F(\alpha')(\xi'/\alpha').$$ \hspace{1cm} (89)

"This is an ordinary differential equation for $(\xi'/\alpha')$, considered as a function of the $\xi''$s". The eigenvalues of this differential equation are the diagonal elements of the matrix representing the dynamical variable $F$.

If one takes the $\xi_r$'s and $\eta_r$'s to be the ordinary $q_r$'s and $p_r$'s of the system at some specified time, and takes $F$ to be the Hamiltonian, $F(\alpha')$ gives the energy of a stationary state of the system, $\alpha'$ being the set of quantum numbers specifying the stationary state. In this case, therefore, Eq. (89) "is just Schrödinger's wave equation". Thus, Dirac states that "the eigenfunctions of Schrödinger's wave equation are just the transformation functions that enable one to transform from the $q$-scheme of matrix representation to a scheme in which the Hamiltonian is a diagonal matrix".

Getting the transformation function $(\xi'/\alpha')$ from Eq. (89), one can calculate in accordance with Eq. (88a) the $\alpha$-scheme of matrix representation of the

\(^4\text{42) It is sufficient to prove that if the theorem is true for any two functions } f_1 \text{ and } f_2, \text{ it is true for } f_1 + f_2 \text{ and } f_1f_2.\)
dynamical variable \( f(\xi_r, \eta_r) \) from the formula

\[
f(\xi_r, \eta_r)(\alpha' \alpha'') = \int \left( \alpha'/\xi' \right) d\xi' f \left( \xi'_r, -i\hbar \frac{\partial}{\partial \xi'_r} \right) \left( \xi'/\alpha'' \right).
\]

On taking \((q_r, p_r)\) for \((\xi_r, \eta_r)\) and the Hamiltonian \(H(q_r, P_r)\) for \(F\), this formula given by Dirac corresponds to Eq. (66) in Sec. 3, §1, Chap. 2 of the present volume, which was given by Schrödinger in his “paper on the relationship”.

Schrödinger proved that, expressed in Dirac’s notation, \( f(\xi_r, \eta_r)(\alpha' \alpha'') \) calculated from the wave functions \((\xi' / \alpha')\) and \((\alpha' / \xi')\) according to this formula is equivalent, on account of the completeness of the wave functions, to the matrix element of the relevant dynamical variable in matrix mechanics. Dirac, on the other hand, takes this formula to be the transformation from the \(\xi\)-scheme to the \(\alpha\)-scheme of matrix representation by introducing the \((\xi' \xi'')\) elements with continuous range of \(\xi''\)’s besides the \((\alpha' \alpha'')\) elements with discrete set of \(\alpha''\)’s, in order to use matrix mechanics to treat collision problems, which were successfully treated in a quantum-theoretical way for the first time by Born with the use of wave mechanics.

Matrix mechanics originally premised phenomenologically the existence of matrix elements which are labeled by discrete quantum numbers. Concerning this Schrödinger proposed the substantialistic wave mechanics including only “demands in which no word occurs of ‘integers’ any more”, and proved the existence of matrix elements labeled by discrete quantum numbers by noting that the wave equation has a discrete set of eigenfunctions. Dirac introduces in addition matrix elements labeled by continuous parameters, still premising those labeled by discrete parameters, and shows that the transformation functions from a continuous scheme of matrix representation to a discrete one are just the eigenfunctions of the Schrödinger equation.

Thus, this procedure of Dirac assures, in the framework of matrix mechanics, the existence of matrix elements labeled by discrete sets of quantum numbers. This, together with the interpretation of transformation function or wave function as transition amplitude, is an important moment of the development of the phenomenological matrix mechanics into the essentialistic quantum mechanics by the mediation of the substantialistic wave mechanics.

Now, for systems in which the Hamiltonian \(H\) involves the time \(t\) explicitly, there will be no matrix scheme with respect to which \(H\) is diagonal because there is no set of constants of integration that do not involve \(t\) explicitly. For such cases, Dirac takes the values of \(q_r\) and \(H_r\) of \(q\) and \(H\) at the time \(t = \tau\).
respectively, and denotes by $\alpha$ “a set of constants of integration that can be expressed as functions of the $q$'s, $p$'s and $t$ at the arbitrary time $t$ that do not involve the parameter $\tau$”. Namely, he considers a set of constants of integration that are obtained separately at every time. From the condition that the $\alpha$’s must satisfy, the $(\alpha'\alpha'')$ element of $df/d\tau$, i.e., $(df/d\tau)(\alpha'\alpha'')$, is seen to be equal to the differential coefficient of $f(\alpha'\alpha'')$, i.e., $\partial f(\alpha'\alpha'')/\partial \tau$, where $f$ is any function of the $q$'s and $p$'s.

Further, if $f$ does not involve $\tau$ explicitly, one has $(fH_\tau - H_\tau f) = i\hbar df/d\tau$. From these relations, $(fH_\tau - H_\tau f)(q'_\alpha')$ can be expressed as the difference between a term involving $f$ in the $q_\tau$-scheme of representation and one involving $f$ in the $\alpha$-scheme of representation, with the use of the transformation function $(q'_\tau/\alpha')$ and integration by parts. On the other hand, the same matrix element can be calculated directly by applying the multiplication rule of matrices and expressed as a difference of the same kind. By comparison of the two expressions Dirac derives the relation

$$H_\tau(q'_\tau/\alpha') = i\hbar \left( \frac{\partial}{\partial \tau} \right) (q'_\tau/\alpha'). \quad (90)$$

If one writes $t$ for $\tau$ and $q$ for $q_\tau$, this becomes identical with Eq. (86a) in which the pair of canonical variables $(\xi_\tau, \eta_\tau)$ is taken to be $(t, -H)$. Thus one gets, from Eq. (90) and Eq. (88a) in which the pair $(\xi_\tau, \eta_\tau)$ is taken to be the pair $(q_\tau, p_\tau)$,

$$H(q_\tau, p_\tau)(q'\alpha') = H \left( q_\tau, -i\hbar \frac{\partial}{\partial q_\tau} \right) (q'/\alpha') = i\hbar \left( \frac{\partial}{\partial t} \right) (q'/\alpha'), \quad (91)$$

“which is Schrödinger’s wave equation for Hamiltonians that involve the time explicitly”.

Dirac relates in this wavy the transformation function to the wave function. Born’s result still remains, however, to be related to matrix mechanics. Thus, “to obtain physical results from the matrix theory”, Dirac introduces newly “the only assumption one needs make”. Namely, he assumes that “the diagonal elements of a matrix, whose rows and columns refer to the $\xi$’s say representing a constant of integration, $g$ say, of the dynamical system, determine the average values of the function $g(\xi_\tau, \eta_\tau)$ over the whole of $\eta$-space for each particular set of numerical values for the $\xi$’s”. Dirac says that this is “in the same way in which they certainly would in the limiting case of large quantum numbers".
Thus, if

\[ g(\xi'\xi'') = g(\xi')\delta(\xi' - \xi'') , \]

(92)

when the \(\xi''\)'s are nearly equal to \(\xi'\)'s, \(g(\xi')\) is assumed to be the average value of \(g\) over the whole of \(\eta\)-space when \(\xi_r = \xi'_r\). In the case when the diagonal elements of the matrix \(g\) are finite without requiring the removal of the factor \(\delta(\xi' - \xi'')\), "the corresponding assumption is that the diagonal element \(g(\xi'\xi')\) is equal to \((2\pi\hbar)^{-1}\) times the integral of \(g(\xi_r, \eta_r)\) over all \(\eta\)-space when \(\xi_r = \xi'_r\).”

Letting \(g_1, g_2, \ldots, g_u\) be a set of \(g\)'s that commute with one another and are independent functions of the \(\xi\)'s and \(\eta\)'s, Dirac proceeds to find the matrix that represents

\[ \delta(g_1 - g'_1)\delta(g_2 - g'_2) \cdots \delta(g_u - g'_u) = \delta(g - g') . \]

(93)

The integral of this with respect to the parameter \(g'\),

\[ \int_{g'}^{g''} \delta(g - g')dg' , \]

will be the matrix which represents the function of \(g\)'s that equals unity for \(g'_r < g_r < g''_r\) and zero otherwise. The diagonal elements of this matrix will then give the \(\eta\)-average of this function, which is just the fraction of the whole of \(\eta\)-space for which \(g'_r < g_r < g''_r\).

Because of the properties of the \(\delta\)-function, \(\int \delta(x)dx = 1\) and \(x\delta(x) = 0\), one must have \(\int \delta(g - g')dg' = 1\) and

\[ (g_r - g'_r)\delta(g - g') = \delta(g - g')(g_r - g'_r) = 0 \]

(94)

for each \(r\). Dirac shows that the matrix with elements

\[ \delta(g - g')(\xi\xi') = (\xi'/g')(g'/\xi'') \]

(95)

satisfies these conditions, by using Eq. (92) and "the orthogonal and normalized properties" of the transformation functions \((\xi'/g')\) and \((g'/\xi'')\). Therefore, the volume of \(\eta\)-space for which the \(g\)'s lie between their given numerical values

\[ \int d\xi_r d\eta_r . \]

43) "In the limiting case of large quantum numbers", as Dirac says, the quantum condition in the old quantum mechanics is valid, and \(\hbar (= 2\pi\hbar)\) serves as the fundamental unit of the phase integral \(\int\! d\xi_r d\eta_r\).
when $\xi_r = \xi_r's$ is given by

$$
(\xi'/g')dg'(g'/\xi') = |(\xi'/g')|^2 dg'.
$$

(96)

This expression of Dirac forms, as is seen from Dirac's words quoted at the end of §3 of the present section, an extension of Heisenberg's idea that $|S_{\alpha\beta}|^2$ is the probability as described also in the same §3.

To derive Born's result, it is sufficient to find "the transformation function that connects the final components of momentum of the electron, $p_F$, with the initial components $p_I"$. $|p'_F/p'_I|^2 dp_F$ is then the probability that the electron will be scattered with a momentum lying in the range $dp_F$. If the coordinates of the electron at the time $t$ are $x_t$, one has

$$
(\xi'_t/p'_I) = \int (x'_t/p'_F) dp'_F (p'_F/p'_I),
$$

(97)

where the transformation functions $(x'_t/p'_I)$ and $(x'_t/p'_F)$ are the wave functions appropriate to the incident and emerging electrons respectively. This equation represents the resolution of the former wave into the latter waves, the amplitudes of the various components being $(p'_F/p'_I)$.

In this way Dirac proved Born's result from "the only assumption one needs make". Because Eq. (97) becomes $\psi = \sum_n c_n \psi_n$ in the notation used by Born, with $|c_n|^2$ giving the probability, it would suffice to assume directly that $|p'_F/p'_I|^2 dp'_F$ is the probability. Dirac introduced "the only assumption one needs make" in order to formulate collision problems in accordance with the way of thinking in "the matrix theory".

Dirac showed that Schrödinger's eigenfunctions mean the transformation functions from the $q$-scheme of matrix representation to the scheme in which the Hamiltonian is diagonal. On the other hand, London showed previously, as has been described in §2 of the present section, Schrödinger's eigenfunctions mean the axes of coordinates spanning a Hilbert function space. These results imply that axes of coordinates of a certain kind are related to schemes of matrix representation. This problem will be discussed in §7 of the present section.

§6. **Formulation based on the concept of probability amplitude**

Jordan, independently of Dirac, developed a theory substantially identical with Dirac's in his paper "On a New Establishment of Quantum
The Unification as Quantum Mechanics received on Dec. 18, 1926, half a month later than the Dirac paper we have just mentioned. Jordan stated in a footnote added in proof that "being considered from a somewhat different side essentially the same facts as are explained in the present work are found, in a forthcoming work of Mr. P. A. M. Dirac". Jordan developed his theory, however, on the "postulate" that the wave function should be the probability amplitude, in contrast to Dirac. In a footnote Jordan cited London's paper we treated in §2 of the present section, saying that partially the same result as Jordan's was independently obtained by London.

According to Jordan, in "a paper in press" Pauli "has proposed, in connection with Born's considerations, the physical interpretation" that, if \( \varphi_n(q) \) is a normalized eigenfunction of Schrödinger's wave equation, \( |\varphi_n(q)|^2dq \) gives the probability that, when the system is in the state \( n \), the coordinate \( q \) has a value in the interval \( q, q + dq \). Born interpreted that \( |c_n(t)|^2 \), given by the expansion \( \sum_n c_n(t)\psi_n(q) \) of a wave function, should give the probability for the system to be in the state \( n \). Jordan says that Pauli's interpretation "is closely related to Born's interpretation" and that "both interpretations are physically so immediately plausible and natural that any detailed explanation seems unnecessary". Jordan says also that Pauli "considers the generalization" that \( |\varphi(q, \beta_0)|^2dq \) measures the probability", when a continuously varying dynamical quantity \( \beta \) takes the value \( \beta_0 \), and "the function \( \varphi(q, \beta) \) is called by Pauli the probability amplitude".

Jordan assumes, on the basis of Pauli's suggestion, that "about two mechanical quantities \( q \) and \( \beta \) which stand by each other in a completely determined relationship, there exist functions \( \varphi(x, y) \) and \( \psi(x, y) \) such that

\[ \varphi(x, y)\psi^*(x, y)dx \]

(98)

is the relative or absolute probability that in the case of \( \beta = y \), the value of \( q \) lies between \( x \) and \( x + dx \), and calls this "Postulate A". He names \( \varphi(x, y) \) "probability amplitude" and \( \psi(x, y) \) "completion amplitude". (\( \psi^* \) etc. denote the complex conjugates of \( \psi \) etc.) Because Jordan's way of denotation is not

---


45) Pauli's paper in which the spin matrices were introduced. This paper will be treated in §4 of the following section.

46) M. Born, Zeit. f. Phys. 38 (1926), 803; 40 (1926), 167 [same as Refs. 19) and 20) in the present chapter].
easy to read, we shall in the following apply Dirac’s way of denotation and write expression (98) as \( \varphi(q', \beta')\varphi^*(q', \beta')dq' \).

As “Postulate B”, Jordan assumes that “the corresponding functions \( \varphi(\beta', q') \) and \( \bar{\varphi}(\beta', q') \) for the exchanged pair \( \beta \) and \( q \) are given by”

\[
\varphi(\beta', q') = \psi^*(q', \beta') , \quad \bar{\varphi}(\beta', q') = \varphi^*(q', \beta') .
\]

(99)

In “Postulate C”, it is assumed that “the probabilities combine interentially”. Namely, the probability amplitudes are assumed to obey the usual laws of probability, so that \( \varphi_1 + \varphi_2 \) gives the probability amplitude for a pair of exclusive events and \( \varphi_1\varphi_2 \) that for a pair of independent events.

It results from this, as the first conclusion, that if \( \Phi, \chi \) and \( \varphi \) are the probability amplitudes for the pairs of dynamical quantities \( (Q, \beta), (Q, q) \) and \( (q, \beta) \) respectively, then

\[
\Phi(Q', \beta') = \int \chi(Q', q')\varphi(q', \beta')dq' .
\]

(100)

This equation of Jordan gives the transformation of \( \varphi(q', \beta') \) into \( \Phi(Q', \beta') \) through the mediation of \( \chi(Q', q') \). As we shall see later, Jordan connects the probability amplitudes with matrix elements for the reason that this equation satisfies the law of multiplication of matrices with continuous rows and columns.

For \( Q = \beta \) in particular, one has from this equation and Postulate B

\[
\Phi(\beta'', \beta') = \int \bar{\varphi}(\beta'', q')\varphi(q', \beta')dq' .
\]

(101)

Since this should give the probability that \( \beta \) takes the value \( \beta' \) when \( \beta = \beta'' \), one has

\[
\int \bar{\varphi}(\beta'', q')\varphi(q', \beta')dq' = \delta_{\beta'\beta} ,
\]

(102)

if the amplitudes concerned have been properly normalized. In the case of a continuously varying \( \beta \), “this equation is not considered to be a very correct way of writing”.\(^{47}\) Also, if the domain of the values of \( q \) includes

\(^{47}\) Dirac’s \( \delta \)-function described in §5 of the present section gives the appropriate way of expression. Jordan mentions, in another part of his paper under consideration, that \( (\beta'' - \beta')\Phi(\beta'', \beta') = 0 \) and \( (\partial/\partial \beta'' + \partial/\partial \beta')\Phi(\beta'', \beta') = 0 \), for the reason that \( \Phi(\beta'', \beta') = 0 \) except for \( \beta'' = \beta \) and is a function only of the difference \( (\beta'' - \beta') \). \( \delta(\beta'' - \beta') \) has these properties.
a set of discrete points, a summation "should be supplemented" to the integration.

If the probability amplitude for a pair of dynamical quantities \((p, q)\) is given by

\[
\rho(p', q') = \exp \left( -\frac{ip'q'}{\hbar} \right),
\]  

(103)

Jordan "calls \(p\) the momentum canonically conjugate to \(q\)."\(^{48}\) From "the orthogonality relation" of complex exponential functions, the completion amplitude of \(\rho(p', q')\) is also given by \(\exp(-ip'q'/\hbar)\). There results thus "the law that for a given value of \(q\) all possible values of \(p\) are equally probable". It is assumed that "for every \(q\) there exists a conjugate momentum" ("Postulate D"). \(\rho(p'q')\) satisfies the equations

\[
\left( p' - i\hbar \frac{\partial}{\partial q'} \right) \rho(p', q') = 0,
\]  

(104a)

\[
\left( i\hbar \frac{\partial}{\partial q'} - p' \right) \rho(p', q') = 0.
\]  

(104b)

With the use of these equations, Jordan argues that the operator \(-i\hbar \partial/\partial q\) "corresponds" to \(p\), and the commutation relation \([p, q]=pq - qp = -i\hbar\) "holds".

Then, in order to show that the formula for the multiplication of the probability amplitudes \(\varphi(q', \beta')\), \(\chi(Q', q')\) and \(\Phi(Q', \beta')\), Eq. (100), can be derived from equations of the Schrödinger type, Jordan considers the transformation among the sets of canonical variables \((p, q)\), \((\alpha, \beta)\) and \((P, Q)\). For the transformation from \((p, q)\) to \((\alpha, \beta)\),

\[
\alpha = f(p, q) = TpT^{-1}, \quad \beta = g(p, q) = TqT^{-1},
\]  

(105)

using the operator \(T = T(p, q) = T(-i\hbar \partial/\partial q, q)\), it can be shown that \([\alpha, \beta] = -i\hbar.\(^{49}\) The probability amplitude \(\varphi(q', \beta')\) is assumed to satisfy the differential equations

\[
\left\{ f \left( -i\hbar \frac{\partial}{\partial q'}, q' \right) - i\hbar \frac{\partial}{\partial \beta'} \right\} \varphi(q', \beta') = 0,
\]  

(106a)

\[
\left\{ g \left( -i\hbar \frac{\partial}{\partial q'}, q' \right) - \beta' \right\} \varphi(q', \beta') = 0.
\]  

(106b)

\(^{48}\)Jordan uses the notation \(\varepsilon\) for \(\hbar/2\pi i\).

\(^{49}\)P. Jordan, Zeit. f. Phys. 38 (1926), 513 [same as Ref. 29] in §2 of the present section.
In particular, if Eq. (105) is written as

\[ \alpha = f(p, q) = -q, \quad \beta = g(p, q) = p, \quad (105') \]

Eq. (106) becomes

\[ \left( q' + i\hbar \frac{\partial}{\partial p'} \right) \varphi(q', p') = 0, \quad (106'a) \]
\[ \left( i\hbar \frac{\partial}{\partial q'} + p' \right) \varphi(q', p') = 0. \quad (106'b) \]

These equations have the solution \( \varphi(q', p') = \exp(iq'p'/\hbar) \). Because \( \varphi^*(q', p') = \varphi(p', q')\), it is seen that \( \varphi(p', q') \) is equal to \( \rho(p', q') \) given by Eq. (103). This shows that Eq. (106) gives Postulate D.

Further, taking the transformation between \((p, q)\) and \((P, Q)\)

\[ p = p(P, Q) = SPS^{-1}, \quad q = q(P, Q) = SQS^{-1}, \quad (107) \]

by means of \( S = S(P, Q) = S(-i\hbar \partial / \partial Q, Q) \), and putting \( f(p, q) = F(P, Q) \) and \( g(p, q) = G(P, Q) \) in Eq. (105), one gets for the probability amplitude \( \Phi(Q', \beta') \) the differential equations corresponding to Eq. (106)

\[ \left[ F \left( -i\hbar \frac{\partial}{\partial Q'}, Q' \right) - i\hbar \frac{\partial}{\partial \beta'} \right] \Phi(Q', \beta') = 0, \quad (108a) \]
\[ \left[ G \left( -i\hbar \frac{\partial}{\partial Q'}, Q' \right) - \beta' \right] \Phi(Q', \beta') = 0. \quad (108b) \]

If one uses for \( \phi(q', \beta') \) in Eq. (106)

\[ \Phi(Q', \beta') = S\varphi(q', \beta'), \quad (109) \]

\( \Phi(Q', \beta') \) given by this equation satisfies Eq. (108).\footnote{The operator \( S \) transforms the function \( \varphi(q', \beta') \) at the same time as it transforms the variable from \( q' \) to \( Q' \).}

Then, in order to solve Eq. (106), Jordan takes

\[ \varphi(q', \beta') = \int \Phi(Q', \beta') \chi^*(Q', q') dQ' \quad (110) \]
“by generalizing the method of the Laplace transformation”. He shows that
\( \chi(Q',q') \) given here satisfies the equations

\[
\begin{align*}
\left\{ p \left( -i\hbar \frac{\partial}{\partial Q'}, Q' \right) - i\hbar \frac{\partial}{\partial q'} \right\} \chi(Q', q') &= 0, \\
\left\{ q \left( -i\hbar \frac{\partial}{\partial Q'}, Q' \right) - q' \right\} \chi(Q', q') &= 0,
\end{align*}
\]

and therefore is a probability amplitude.

From Eqs. (109) and (110) it is seen that

\[
S^{-1} = \int dQ' \chi^*(Q', q') \cdots.
\]

Also, starting from the adjoint differential equations of Eq. (106),\(^{51}\) one can show by a similar procedure that

\[
S^{-1} = \int dq' \chi(Q', q') \cdots. \tag{112'}
\]

Therefore, with the use of the relation \( \chi^*(Q', q') = \bar{\chi}(q', Q') \), one gets

\[
SS^{-1} = \int dq' dQ' \chi(Q'', q') \bar{\chi}(q', Q') \cdots, \tag{113}
\]

which means that \( SS^{-1} \) is the unit operator on account of Eq. (102).

Equations (109) and (112') give

\[
\Phi(Q', \beta') = \int dq' \chi(Q', q') \varphi(q', \beta'). \tag{114}
\]

This is identical with the formula (100) for the “product” of probability amplitudes. Jordan concludes that “we have thus confirmed the property of the amplitudes demanded by Postulate C”. Namely, it is shown by Jordan that the differential equations (106) assumed for the probability amplitude are equivalent to Postulate C and Postulate D.

Next, Jordan introduces matrices with continuous rows and columns, by saying that “if \( M_1(x,y) \) and \( M_2(x,y) \) are two matrices, one usually calls

\(^{51}\)Let \( L[u] = 0 \) be a differential equation with variable coefficients. If there is such a differential operator \( \bar{L} \) that \( \nu L[u] - u \bar{L}[\nu] \) can be expressed as the divergence of a vector quantity, \( \bar{L}[u] = 0 \) is the adjoint differential equation of \( L[u] = 0 \).
the matrix
\[ \int M_1(x, z)M_2(z, y)dz \]
the product \( M_1M_2 \). Because Eq. (110) or (114) is seen to be of this form, he "calls a probability amplitude a matrix".

On the other hand, "Schrödinger's wave equation in which the time is eliminated", that is, the time-dependent Schrödinger equation, is obtained as the "special case" of Eq. (106a) in which the pair of canonical variables \( \alpha \) and \( \beta \) is taken to be that of the Hamiltonian and the time. In this case, the probability amplitude \( \varphi(q', t) \) is identical with the wave function. Also, if \( W \) is the energy of the system, \( \varphi(q', W') \) becomes the time-independent wave function. According to Schrödinger, the matrix \( A(W', W'') \) of a dynamical quantity \( A(p, q) \) is given by

\[ A(W', W'') = \int \bar{\psi}(W', q')A(-i\hbar \frac{\partial}{\partial q'}, q') \varphi(q', W'')dq', \quad (115) \]

where \( \bar{\psi}(W', q') = \varphi^*(q', W') \).\(^{52}\) If \( A(-i\hbar \frac{\partial}{\partial q'}, q') \) is considered an integral transformation in terms of the function \( a(q', q'') \), similar to \( S \) being the canonical transformation discussed above, one has

\[ A(W', W'') = \int dq'dq'' \bar{\psi}(W', q')a(q', q'')\varphi(q'', W''). \quad (115') \]

This is "nothing but Lanczos's assignment" of matrices.

If Eq. (115) is applied to a dynamical quantity \( A(W, t) \), one gets

\[ A(W', W'') = \int \bar{\psi}(W', t')A(-i\hbar \frac{\partial}{\partial t'}, t') \varphi(t', W'')dt'. \quad (116) \]

Further, if \( \rho(t', W') \) given by Eq. (103) is used for the probability amplitude \( \psi(t', W') \) in this equation, one has

\[ A(W', W'') = \int e^{-iW't'/\hbar}A(-i\hbar \frac{\partial}{\partial t'}, t') e^{iW''t'/\hbar}dt'. \quad (116') \]

This gives "Born-Wiener's assignment" of matrices.

After these considerations, Jordan concluded that "the forms developed up to now of quantum mechanics are contained as special cases in a general formal

\(^{52}\)Here also, following Dirac, we denote matrices as \( A(W'W'') \) to distinguish them from functions \( A(p, q) \). Jordan uses the notation \( A(W, V) \) instead of \( A(W'W'') \).
theory”. Starting from the concept of the wave function being a probability amplitude, on the basis of wave mechanics, Jordan arrived thus at a theory which is essentially equivalent to Dirac’s theory, by applying the method of canonical transformations used in matrix mechanics to wave mechanics. Jordan’s probability amplitude \( \varphi(q', \beta') \) and the complex conjugate of its completion amplitude \( \psi^*(\beta', q') \) correspond to Dirac’s transformation functions \( (q'/\beta') \) and \( (\beta'/q') \) respectively, satisfying relations equivalent to those obeyed by the latter functions. It may be said that Jordan’s theory was lacking in the introduction of the \( \delta \)-function and the easily understandable notations, in comparison with Dirac’s theory.

Dirac’s theory and Jordan’s theory showed that, one can arrive at one and the same quantum mechanics on the basis of either matrix mechanics or wave mechanics, if one generalizes the wave function \( \varphi(q', t') \) as a function of the space coordinate \( q \) and time \( t \) or the wave function \( \varphi(q', W) \) as a function of \( q \) and the energy \( W \) to the concept of the probability amplitude \( \varphi(q', \beta') \), and connects it with the transformation function, or the operator, of the canonical transformation from the pair of the canonical variables \( p \) and \( q \) to that of the canonical variables \( \alpha \) and \( \beta \).

§7. The unified formulation with the use of abstract Hilbert space

Dirac’s theory and Jordan’s theory were arranged and systematized by mathematicians Hilbert, v. Neumann and Nordheim. In their paper “On the Foundation of Quantum Mechanics”, \(^{53}\) received on Apr. 4, 1927, Hilbert, v. Neumann and Nordheim say that “Jordan has given in connection with an idea by Pauli a formulation and foundation of the theory, which is of greater generality and seems to be adapted very well to the physical character of the underlying phenomena”, and that “independently of Jordan, Dirac is led to similar views”.

This paper by Hilbert et al. “is evolved from a lecture which D. Hilbert held in the winter semester 1926/27 on the recent development of quantum mechanics, and the preparation of which is done with essential help by L. Nordheim”. \(^{54}\) “Important pieces of the mathematical implementation come from


\(^{54}\)Hilbert’s lecture was held at Göttingen University. The papers by Jordan and by Born, whom Jordan thanked “for a bit of friendly advice and stimulation”, were worked out at Göttingen University.
J. v. Neumann”, and “in the final writing of the treatise L. Nordheim is concerned”. Hilbert et al. “believe that the formulation of Jordan’s and Dirac’s thought, to which we are here led, has become considerably simpler and therefore more transparently and easily understandable”.

Hilbert-v. Neumann-Nordheim express Jordan’s probability amplitude of a pair of mechanical quantities $F_1(pq)$ and $F_2(pq)$ as $\phi(xy, F_1F_2)$, $x$ and $y$ being the values taken by $F_1$ and $F_2$ respectively. Jordan expressed it simply as $\phi(x,y)$. Since this notation by Jordan is not easy to read, in the foregoing §6 we have written it as $\phi(F_1^*F_2^*)$ following to Dirac. The formula for multiplication of probability amplitudes in Jordan’s saying is now expressed as

$$\phi(xz; F_1F_3) = \int \phi(xy; F_1F_2)\phi(yz; F_2F_3)dy.$$  \hfill (117)

Because $S$ in Eq. (109) as well as $T$ in Eq. (105) of a similar kind used by Jordan are operators that transform functional forms and independent variables at the same time, Hilbert et al. introduce notations such as $S^{(\tau)}_\varphi$ and $T^{(\tau)}_\varphi$. The operator $T^{(\tau)}_\varphi$ is defined to transform a function $f(y)$ of $y$ according to the formula

$$T^{(\tau)}_\varphi(f(y)) = \int \varphi(xy)f(y)dy,$$  \hfill (118)

where the function $\varphi(xy)$ is called “the kernel of the operator $T^{(\tau)}_\varphi$”. It is also written as $T^{(\tau)}_{\varphi}f(y)$ when it is necessary to indicate its kernel explicitly. One has thus

$$T^{(\tau)}_{\varphi}T^{(\tau)}_\psi = T^{(\tau)}_{\chi},$$  \hfill (119)

“The unit operator” $1^{(\tau)}_\varphi$ is defined in terms of the $\delta$-function as

$$1^{(\tau)}_\varphi f(y) = f(x) = \int \delta(x-y)f(y)dy.$$  \hfill (120)

Because

$$T^{(\tau)}_\varphi \delta(u-y) = \int \varphi(xu)\delta(u-y)du = \varphi(xy),$$  \hfill (121)

$\varphi(xy)$ giving the kernel of $T^{(\tau)}_\varphi$ is equal to the result of $\delta(u-y)$ operated on by $T^{(\tau)}_\varphi$. If $(T^{(\tau)}_\varphi)^{-1}$, the inverse operator of $T^{(\tau)}_\varphi$, is expressed as $T^{-(\tau)}_{\varphi^{-1}} = T^{(\tau)}_{\varphi^{-1}}$, ...
its kernel $\varphi^{-1}$ satisfies the relation
\[ \int \varphi(xu)\varphi^{-1}(uy)du = \delta(x - y). \] (122)

Because $\delta(x - y)$ satisfies the formulae
\[ \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) \delta(x - y) = 0, \quad (x - y)\delta(x - y) = 0, \] (123)

one obtains, by operating $T_{\varphi}^{-1}(x)$ on these formulae and making use of the relation $T_{\varphi}^{-1}(x)T_{\varphi}^{-1}(y) = 1(x)$ and Eq. (121), the relations
\[ \{ T_{\varphi}^{-1}(x)T_{\varphi}^{-1}(y) \} \varphi(xy) = 0, \] (124a)
\[ \{ T_{\varphi}^{-1}(y) - y \} \varphi(xy) = 0. \] (124b)

If one defines $F^{(x)}$ and $G^{(x)}$ to be
\[ F^{(x)} = T_{\varphi}^{-1}(x) - i\hbar \frac{\partial}{\partial x} T_{\varphi}^{-1}(x), \] (125a)
\[ G^{(x)} = T_{\varphi}^{-1}(y) - y, \] (125b)

one has the functional equations
\[ \left\{ F^{(x)} - i\hbar \frac{\partial}{\partial y} \right\} \varphi(xy) = 0, \] (126a)
\[ \{ G^{(x)} - y \} \varphi(xy) = 0. \] (126b)

These correspond to Eqs. (106a) and (106b) given by Jordan. If the space coordinate $q$, the Hamiltonian $H(p, q)$ and the energy $W$ are taken for $F_1, F_2$ and $y$ respectively, $\varphi(xy) = \varphi(xW, qH)$ gives Schrödinger's wave function.

In this way, Hilbert-v. Neumann-Nordheim made the representation of Jordan's theory "easily understandable", with the use of Dirac's $\delta$-function.\(^{55}\)

\(^{55}\)Citing Dirac's and Hilbert-v. Neumann-Nordheim's papers and making use of the $\delta$-function, Jordan also gave, in his paper "On a New Establishment of Quantum Mechanics II" received on Jun. 3, 1927 [Zeit. f. Phys. 44 (1927), 1], "a logical generalization — and at the same time an essential simplification — " of the representation of his theory given in his paper I.
v. Neumann, in his paper "Mathematical Establishment of the Quantum Mechanics" received on May 20, 1927, showed that wave mechanics of which the wave functions are grasped as probability amplitude, on one hand, and matrix mechanics of which transformation functions (matrix elements of canonical transformation) are grasped as wave functions, on the other hand, can be formulated in unification by means of the linear algebra in the abstract Hilbert space. Since the results arrived at by v. Neumann in this paper are adopted in present-day text books on quantum mechanics and are well known, we shall in the following treat mainly the ideas of his formulation.

In "the matrix formulation", one considers "the energy matrix"

$$H = \{h_{\mu\nu}\} \ (\mu, \nu = 1, 2, \ldots)$$  \hspace{1cm} (127)

and "the transformation matrix"

$$S = \{s_{\mu\nu}\} \ (\mu, \nu = 1, 2, \ldots),$$  \hspace{1cm} (128)

which makes $S^{-1}HS = W$ diagonal. $H$ is assumed to be a Hermitian matrix ($h_{\mu\nu} = \bar{h}_{\nu\mu}$), and $S$ is a "gonalortho" matrix ($\sum_{\rho} s_{\mu\rho} \bar{s}_{\nu\rho} = \delta_{\mu\nu}$), where $h_{\mu,\nu}, s_{\nu\rho}$ denote the complex conjugates of $h_{\mu\nu}, s_{\nu\rho}$.

If "the diagonal terms" of $W$, that is, its diagonal elements are denoted by $w_1, w_2, \ldots$, one gets from $HS = SW$

$$\sum_{\rho=1}^{\infty} h_{\mu\rho} s_{\rho\nu} = S_{\mu\nu} w_{\nu}.$$  \hspace{1cm} (129)

Namely, "every column of $S$ is therefore a solution of the eigenvalue problem of finding such a sequence $x_1, x_2, \ldots$ that is transformed by $H$ into a multiple — into $w$-times for example — of itself". This means the same thing as the transformation into principal axes used by Born-Heisenberg-Jordan, which we have treated in Sec. 2, §4, Chap. 1 of the present volume.

"In the case of the differential equation formulation the situation is still clearer: the eigenvalue problem is given from the beginning". There is a differential operator $H$, and one searches for a function $\psi$ (as a function of space

---

56) J. v. Neumann, Göttinger Nachrichten 1927, 1. The contents of this paper and his paper on quantum mechanical measurements, which we shall consider in the following chapter, are included in his famous book "Mathematische Grundlagen der Quantenmechanik" (Springer, 1932).
coordinates) for which

\[ H\psi = w\psi, \]

that is, "a function which is transformed by \( H \) into a multiple of itself". v. Neumann asks next "What is the common essential feature of all these cases?" "Clearly the one, that in every case there is a manifold of given quantities and a linear operator \( H \) in the manifold". "Every time for all the (real) values \( w \) it is looked for whether there is a nonvanishing element \( f \) of this manifold, such that \( Hf = wf \) is valid". v. Neumann says that "it is now our task, by this unified formulation to come to a unified problem", and "we shall carry out this, by proving that both the manifolds just quoted are essentially identical with each other, that is, they can both be produced from a single manifold by mere renaming".

This single manifold is an "abstract Hilbert space \( \mathcal{H} \)". \( \mathcal{H} \) is a space of infinitely many dimensions. For arbitrary elements \( f, g, \ldots \) in it, there should exist in it also an addition \( f + g \) and a multiplication \( af \), \( a \) being a complex number. Furthermore there should exist a "Hermite-symmetric bilinear form" \( Q(f, g) \) with the following properties: 1) \( Q(af, g) = aQ(f, g) \), 2) \( Q(f_1 + f_2, g) = Q(f_1, g) + Q(f_2, g) \), 3) \( Q(f, g) = Q(g, f) \), 4) \( Q(f) \equiv Q(f, f) \geq 0 \) (the equality holding only when \( f = 0 \)). \( f, g, \ldots \) "satisfy the well-known calculation rules of the analogous operations on vectors".

In the case of matrix mechanics, \( x_\mu = S_{\mu\nu} (\mu = 1, 2, \ldots) \) give the components of a vector, and one has \( Q(x, y) = \sum_\mu \bar{x}_\mu y_\mu \). This is a "discontinuous realization" of \( \mathcal{H} \).\(^{57}\) In the case of wave mechanics, \( \varphi(q) \) with various \( q \)'s are the components of a vector, and one has \( Q(\varphi, \psi) = \int \bar{\varphi}(q) \psi(q) dq \). This is a "continuous representation" of \( \mathcal{H} \). If one develops \( \varphi(q) \) and \( \psi(q) \) in terms of a complete set of orthonormal functions \( \varphi_\mu(q) (\mu = 1, 2, \ldots) \) with their respective coefficients \( f_\mu \) and \( g_\mu \), one has \( Q(\varphi, \psi) = \sum_{\mu=1}^{\infty} f_\mu \bar{g}_\mu \), getting a "discontinuous representation."\(^{58}\)

\(^{57}\)As we have seen in Sec. 4, §2, Chap. 1 of the present volume, Born and Wiener, considering that "the physical meaning is not clear" of \( x_\mu \), introduced \( x = \sum_\mu x_\mu \exp(i\omega_\mu t/\hbar) \) to present "Born-Wiener's assignment" of matrices in Jordan's saying [cf. Eq. (116') in the preceding subsection, §6].

\(^{58}\)This corresponds to "the Hilbert function space" in the case of London's paper, which we have mentioned in §2 of the present section. \( \varphi_\mu(q) \) with various \( q \)'s are the components of the \( \mu \)th unit vector, and \( f_\mu \) is the inner product of this unit vector with the vector whose components are given by \( \varphi(q) \).
In a "discontinuous realization" of $\hat{H}$ there exist "linear operators". In the case of matrix mechanics, they are expressed as $(Tx)_\mu = \sum_\mu T_{\mu \nu}x_\nu$ by means of a matrix $T = \{T_{\mu \nu}\}$. On the other hand, in the case of "continuous realization" one has to consider expressions of the form $T\psi(q) = \int \Phi(q,q')\psi(q')dq'$. In order to retain the meaning of this form for the unit operator $T = 1$, "one must consider 'improper' integral kernels, as Dirac does". Namely, one has to treat the case in which $\Phi(q,q') = \delta(q - q')$. Because of this, in the case of wave mechanics v. Neumann takes the "discontinuous realization" with the use of a complete set of functions to carry on his study. In this case, operators are expressed as matrices calculated by means of Schrödinger's method.

However, in wave mechanics, there are those cases in which $Q(\varphi, \varphi)$ is not finite so that the eigenvalues make a "continuous spectrum", as is for the eigenstates of momentum or of position. Because the existence of $Q(\varphi, \varphi)$ is a condition for $\hat{H}$, these cases are not included in $\hat{H}$. One should treat them by taking the limit of some kind of process. v. Neumann applies "splitting of the unit" to take such a limit. For example, the development of $\varphi(q)$ in terms of the eigenfunctions $\exp(ikq/\hbar)$ of the momentum operator $p = -i\hbar \partial/\partial q$ corresponds to its Fourier transform

$$\varphi(q) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} e^{-ikq'/\hbar} \varphi(q')dq' \right\} e^{ikq/\hbar}dk . \quad (131)$$

Now, v. Neumann defines the splitting of the unit operator $E(l)$ to be given by

$$E(l)\varphi(q) = \int_{-\infty}^{l} \left\{ \int_{-\infty}^{+\infty} e^{-ikq'/\hbar} \varphi(q')dq' \right\} e^{ikq/\hbar}dk . \quad (132)$$

Namely, $E(l)$ is the operator which cuts off the components with the eigenvalue $k$ greater than $l$ in the spectrum development of the unit operator in terms of the momentum eigenfunctions, involving a kind of "jump". Because of the relation

$$d\{E(l)\varphi(q) = \left\{ \int_{-\infty}^{+\infty} e^{-ilq'/\hbar} \varphi(q')dq' \right\} e^{ilq/\hbar}dl , \quad (133)$$

As we have seen in Dirac's paper taken in §5 of the present section, the eigenfunctions of momentum are given in the form $\varphi(q) = \text{const.} \exp(ipq/\hbar)$, and those of position in the form $\varphi(q) = \delta(q - a)$, $p$ and $a$ being arbitrary real numbers.
one has
\[ \int_{-\infty}^{+\infty} \text{ld}\{E(l)\varphi(q)\} = (-i\hbar) \frac{\partial \varphi(q)}{\partial q}, \] (134)

which is identical with \( p\varphi(q) \).

For the position operator \( q \), the splitting \( E(a) \) of the unit operator into the spectrum of its eigenvalue \( a \) is given by

\[ E(a)\varphi(q) = \begin{cases} \varphi(q) & (a \leq q) \\ 0 & (a > q) \end{cases}, \] (135)

from which it is derived that

\[ \int_{-\infty}^{+\infty} a\text{d}\{E(a)\varphi(q)\} = q\varphi(q). \] (136)

The integrals in Eqs. (134) and (136) are Stieltjes integrals.\(^{60}\) In this way, v. Neumann applies the spectral analysis of the unit operator and the Stieltjes integral, to avoid the \( \delta \)-function.

Generally, v. Neumann considers a complete set of orthonormal vectors \( \varphi_\mu (\mu = 1, 2, \ldots) \) in \( \hat{H} \) and applies the method mentioned above to the equations

\[ f = \sum_{\mu=1}^{\infty} Q(f, \varphi_\mu)\varphi_\mu, \quad Tf = \sum_{\mu=1}^{\infty} Q(f, \varphi_\mu)l_\mu\varphi_\mu, \] (137)

where \( l_\mu \) is the eigenvalue of \( \varphi_\mu \), getting

\[ Tf = \int_{-\infty}^{+\infty} \text{ld}\{E(l)f\} = \int_{-\infty}^{+\infty} \text{ld} \left\{ \sum_{l_\mu \leq l} Q(f, \varphi_\mu)\varphi_\mu \right\}. \] (138)

This expression is then made to tend to the limit of the continuous spectrum of the eigenvalue.

\(^{60}\) A Stieltjes integral \( \int_a^b f(x)d\varphi(x) \) is defined to be \( \lim_{N \to \infty} \sum_{n=1}^{N} f(x'_n)\{\varphi(x_n) - \varphi(x_{n-1})\} \), where \( x_n \)'s are the points at which the interval \([a, b]\) is divided into \( N \) subdivisions, \( x_0 \) and \( x_N \) being equal to \( a \) and \( b \) respectively, and \( x'_n \) is a point in the \( n \)th subdivision. If one brings Eq. (133) into Eq. (134) and interchanges the order of the integrations with respect to \( q' \) and \( l \), one will meet the derivative of Dirac's \( \delta(q' - q) \).
Now, let $\psi_1, \psi_2, \ldots$ be another complete set of orthonormal vectors in $\mathcal{H}$ independent of the set $\varphi_1, \varphi_2, \ldots$. Putting

$$[A] = \sum_{\mu, \nu=1}^{\infty} |Q(\varphi_\mu, A\psi_\nu)|^2$$

(139)

for a Hermitian operator $A$, one has with the use of the notation $Q(f) = Q(f, f)$ the relation

$$[A] = \sum_{\nu=1}^{\infty} Q(A\psi_\nu) = \sum_{\mu=1}^{\infty} Q(A\varphi_\mu),$$

(139')

which shows that $[A]$ is independent of the $\varphi_\mu$'s and $\psi_\mu$'s. Because it is seen that $[A] = \sum_{\mu} \sum_{\nu} |a_{\mu\nu}|^2$ if the matrix representation of $A$ is given by $\{a_{\mu\nu}\}$, v. Neumann calls $\sqrt{[A]}$ "the absolute value" of $A$. He further introduces the projection operator $E(L)$ onto the interval $L$ of the eigenvalue $l$ for which $l' \leq l \leq l''$. $E(L)$ is given as $E(L) = E(l'') - E(l')$ in terms of the splitting of the unit operator $E(1)$.

Then, if one denotes the Hamiltonian operator at time $t$ by $H(t)$, its eigenfunctions by $\varphi_\mu^{(t)}$, a narrow interval around the eigenvalue of $\varphi_\mu^{(t)}$ by $J_\mu$, the projection operator onto $J_\mu$ by $E_t(J_\mu)$, and the corresponding quantities at the time $t_0$ by $H(t_0)$, $\varphi_\nu^{(0)}$'s, $I_\nu$ respectively, one can show by making use of Eq. (139') and the relation $E_t(J_\mu)f = Q(f, \varphi_\mu^{(t)})\varphi_\mu^{(t)}$ that

$$[E_t(J_\mu)E_0(I_\nu)] = Q(Q(\varphi_\nu^{(0)}, \varphi_\mu^{(t)}), \varphi_\mu^{(t)}) = |Q(\varphi_\nu^{(0)}, \varphi_\mu^{(t)})|^2.$$  

(140)

Because $Q(\varphi_\nu^{(0)}, \varphi_\mu^{(t)})$ is the $\nu$th coefficient in the expansion of $\varphi_\mu^{(t)}$ in terms of $\varphi_\nu^{(0)}$'s, v. Neumann concludes that Eq. (140) "is just Born's expression". Namely, it is shown that the absolute value of a probability amplitude is given by "the absolute value" of the projection operator.

In this way, v Neumann showed that matrix mechanics and wave mechanics can be formulated in unification with the use of an abstract Hilbert space. The quantum mechanical states of a system are thus represented by vectors in the abstract Hilbert space. The $\mu$th eigenvector $\varphi_\mu(q)$ of the Schrödinger equation is understood to be the projection of the $\mu$th eigenvector of the Hamiltonian operator onto the eigenvector of the position operator $q$, and is interpreted as an element in the $\mu$th column vector of the matrix for transformation of the principal axes, that is, the matrix of the canonical transformation between these two eigenvectors.
The concept of vectors in the abstract Hilbert space, which represent the quantum mechanical states of a system, is a very important one. These vectors become known later as state vectors. The concept of state vectors, together with that of probability amplitudes grasped as projections between eigenstate vectors, plays essential roles in the problem of measurements in quantum mechanics, which we shall discuss in the next chapter.

3. Developments in Various Problems: Spin Wave Functions, Symmetry for Particle Exchange, Relativistic Formulation

§1. Applications of Born’s statistical interpretation

In Sec. 5, §4, Chap. 2 of the present volume we have considered Born’s theory of collision, in which Schrödinger’s wave function was related for the first time to the probability of finding a particle. Wentzel applied Born’s theory to the Rutherford scattering, in his paper “Two Remarks on the Scattering of Corpuscular Rays as Phenomenon of Diffraction” received on Nov. 19, 1926, about four months later than Born’s paper. Wentzel applied also Born’s statistical interpretation to the photoelectric effect in his paper “On the Theory of Photoelectric Effect” received on the same day as his above-mentioned paper. In these two papers of Wentzel this paper of Born, which was received about one month before Wentzel’s was not yet referred to, although in this paper the time variation of probability amplitude was treated with the use of the time dependent Schrödinger equation.

§1.1 Rutherford scattering

In his paper on the scattering of corpuscular rays, Wentzel says that Born has suggested that $|\psi|^2$ “determines the volume density of mass points”, and uses this hypothesis to find a theory of elastic and inelastic collisions. Wentzel notices also that “Rutherford’s theory of α-ray scattering has found again and again experimental confirmations and even served as the quantitative determination of nuclear charge numbers”. He asks then “Can one show that in certain

---

61) M. Born, Zeit. f. Phys. 37 (1926), 863; 38 (1926), 803 [the same as Refs. 52) and 53) in Chap. 2 of the present volume].
64) M. Born, Zeit. f. Phys. 40 (1926), 167 [the same as Ref. 19) in the present chapter].
limiting cases the statements of the older corpuscular theory and of the new wave theory become identical?" This viewpoint of Wentzel is important, because the Rutherford scattering is, as we have pointed out in the preface to Vol. II of the present book, the one which could be explained without use of \( h \) and served as a key point in the climbing road from the world of classical and substantialistic images to the world of atoms.

Now, Wentzel says that "we can give an affirmative answer to this question in the limiting case of classical mechanics" \(^{65}\) and shows this by assuming that the wave function \( \psi \) can be expressed in the form

\[
\psi = A \exp \left[ \frac{2\pi i S}{h} \right]
\]

in terms of slowly varying functions \( A \) and \( S \). The introduction of this expression into the wave equation and the expansion of it in powers of \( h \) give, "in the first approximation" (in the order of \( h^0 \)) the Hamilton-Jacobi partial differential equation for \( S \), and "in the second approximation" (in the order of \( h^1 \)) the equation

\[
\frac{\partial \log A}{\partial n} = -\frac{\Delta S}{2} |\text{grad} S|,
\]

where \( n \) denotes the distance along "the mechanical orbit" determined by \( \text{grad} S \). On the other hand, Green's theorem applied to a small volume of cross section \( \sigma \) along the orbit gives

\[
\frac{\partial \log(\sigma |\text{grad} S|)}{\partial n} = \frac{\Delta S}{|\text{grad} S|},
\]

and \( |\text{grad} S| \) gives the momentum \( mv \). It results therefore from these two equations that \( A^2 \sigma v = \text{const} \). Because this "guarantees the demand of conservation of particle number" if \( A^2 \) is identified with the volume density of a corpuscular stream, Wentzel considers that this identification "seems to be justified also outside the bounds of validity of classical mechanics" \(^{66}\).

\(^{65}\) As we have seen in Sec. 2, §3, Chap. 2 of the present volume, Wentzel presented, about five months before the paper under consideration, a formulation of the WKB approximation, that is, a semi-classical approximation to the Schrödinger equation.

\(^{66}\) As we have described in Sec. 5, §1, Chap. 2 of the present volume, Schrödinger presented in his fourth paper the time-dependent Schrödinger equation, by which "the continuity equation" is guaranteed not approximately but generally for "the weight function" \( |\psi|^2 \). In Schrödinger's work, however, these equations were not applied to scattering problems.
In order to "treat the scattering of $\alpha$-particles as a phenomenon of diffraction", Wentzel applies Eq. (168) together with Eq. (165) in Sec. 5, §4, Chap. 2 of the present volume, which is due to Born and is called nowadays the Born approximation formula. Wentzel introduces the factor $\exp(-r/R)$ for "the screening by the outer electron shells" into the Coulomb potential between the $\alpha$-particle and nucleus, to avoid the divergence appearing in the integral given by Eq. (165). He takes finally the limit of $R \to \infty$, as is done in present-day textbooks on quantum mechanics, getting the Rutherford scattering formula.

In the Born approximation, use is made only of $u_0$ and $u_1$ appearing in Eq. (163) of the last chapter, which represents the perturbation expansion of the wave function, $\psi(r) = \sum u_n(r)$. Wentzel says that the term of the next order $u_2$ "will be characteristic of wave mechanics, that is, will represent the essential phenomenon of diffraction", but "it remains to be investigated on what condition this would be practically detectable". The fact that the Rutherford scattering formula can be derived exactly from the Schrödinger equation, without the use of the perturbation method about the Coulomb interaction or of the limiting procedure about the screening factor, is proved by Gordon in his paper "On the Collision between Point Charges according to Wave Mechanics",\(^{67}\) received on Mar. 4, 1928, by applying confluent hypergeometric functions to the radial wave functions, in the same way as is described in present-day textbooks on quantum mechanics.

§1.2 Photoelectric effect

Wentzel takes notice of Schrödinger's derivation of the dispersion formula of light, which was carried out in his fourth paper by considering a perturbation periodically varying with time, as we have seen in Sec. 4, §4 in the last chapter. Wentzel says that, when $h\nu$ of the incident light surpasses the ionization energy of the atom, "the highest excited oscillation belongs to the continuous spectrum, and therefore corresponds to photoelectric effects". In this case, "the wave function $\psi$ belonging to the continuous spectrum behaves at large distance $r$ from the nucleus as $r^{-1} \cos(kr - \alpha)$, but such a spherical wave means, in the interpretation given by Born, a stationary electron stream in the radial direction".

\(^{67}\)W. Gordon, Zeit. f. Phys. 48 (1928), 180.
Wentzel writes down the wave function perturbed by the incident light as

$$
\psi = u(x, E_0)e^{2\pi i(mc^2 + E_0)t/\hbar} + \frac{1}{2} \int dE' \frac{a(E_0, E')u(x, E')}{E_0 + \hbar\nu - E'} e^{2\pi i(mc^2 + E_0 + \hbar\nu)t/\hbar},
$$

(142)

according to Schrödinger's fourth paper.\textsuperscript{68} In this expression, \(u(x, E_0)\) and \(u(x, E')\) are the wave functions of the ground and excited states respectively, and \(a(E_0, E')\) is the matrix element of the electric dipole moment determined by these wave functions “according to wave mechanics”. The discrete eigenvalues “are left aside as unimportant”.

Outside the atom, \(u(x, E_0)\) diminishes and in Eq. (142) only the integral term remains. Because of a factor \(r^{-1}\cos(k'r - \alpha)\) in \(u(x, E')\), \(k'\) being given by \(\hbar k'/2\pi = \sqrt{2mE'}\), this integral is divided into two integrals including the outgoing wave \(\exp[i(-k'r + \omega t + \alpha)]\) and the incoming wave \(\exp[i(k'r + \omega t - \alpha)]\). Because their integrands have a singular point at \(E' = E_0 + \hbar\nu\), the contour of integration “is bent a little into the positive imaginary semiplane” at the singular point. Furthermore, the contour of integration is closed by the addition of a curved part of large size in the negative and positive imaginary semiplanes for the outgoing and incoming waves respectively, “according to the well-known procedure”.\textsuperscript{69} There remains consequently only the outgoing wave satisfying Einstein’s law \(E' = E_0 + \hbar\nu\), whose amplitude gives the intensity of photoelectrons when absolutely squared. Though Eq. (142) “seems at first glance to be inconsistent with Einstein’s law”, Wentzel comes to the result just mentioned by slightly bending the contour of integration, in consideration of the fact that “the primary wave is absorbed a little”. Modifications of this kind on contours of integration are often used in present-day textbooks on quantum mechanics.

As we have described in Footnote 61) of the last chapter, Schrödinger solved the time-dependent Schrödinger equation without taking account of the initial condition in the derivation, in his fourth paper on the dispersion formula of light. The intensity of electrons in the photoelectric effect can be calculated directly from the matrix element, if Born’s method to treat the time dependence of the probability amplitude is applied.\textsuperscript{70} Though Wentzel was working from

\textsuperscript{68} Cf. Eq. (121) in the last chapter.

\textsuperscript{69} On these contours of integration the integrands take very small values.

\textsuperscript{70} Cf. Eq. (59) in the present chapter.
the basis of Schrödinger’s method of calculation without the consideration of the initial condition, he made a calculation in which the initial condition was in effect taken into account, with the use of the modification of the contour of integration.

§2. **Relativistic formulation of the theory and the Compton effect**

As has been described in Sec. 1, §§4, 5 of the present chapter, relativistic formulations of the Schrödinger equation were made by Schrödinger in his fourth paper received on Jun. 21, 1926, and by Dirac in his paper received on Aug. 26, 1926. Similar relativistic formulations were also made nearly at the same time by Klein, Fock and Gordon, and were applied to the calculation of the energy levels of the hydrogen atom and of the Compton effect. Dirac formulated relativistically matrix mechanics to treat the Compton effect, too.

§2.1 **Relativistic formulation of the Schrödinger equation**

In a paper received on Apr. 26, 1926, “Quantum Theory and Five-dimensional Relativistic Theory”, 71) Klein makes use of a five-dimensional relativistic theory, which was proposed by Kaluza 72) in an attempt to unify the gravitational and electromagnetic fields, and makes the equation of a charged particle wave-optical, so to speak, deriving a relativistic wave equation. If the wave is sinusoidally varying with a period \( h \) in the direction of the fifth coordinate, the wave obeys in the ordinary four-dimensional space the equation

\[
\left\{ \Delta - \left( \frac{1}{c^2} \right) \frac{\partial^2}{\partial t^2} - \left( \frac{2\pi mc}{h^2} \right)^2 \right\} \psi = 0, \tag{143}
\]

which is the same as the above-mentioned wave equation given by Schrödinger and by Dirac.

Fock derives the same relativistic wave equation in his paper “On Schrödinger’s Wave Mechanics”, 73) received on Jun. 11, 1926, from the variation principle \( \delta \int Qd\tau_1 d\tau_2 \cdots d\tau_n dt = 0 \), where \( d\tau_i = dx_i dy_i dz_i \) and \( Q \) is the quadratic form that is obtained from the relativistic Hamilton-Jacobi partial

---

differential equation by replacing $\partial S/\partial t$ with $-E$ and $\partial S/\partial x_i$ etc. with $-E(\partial \psi/\partial x_i)/(\partial \psi/\partial t)$ etc., and multiplying the result by $(\partial \psi/\partial t)^2$. The replacements used by Fock are substantially identical with the above-mentioned ones used by Schrödinger or Dirac. Fock applies the relativistic wave equation to the calculation of the energy levels of the hydrogen atom. The energy level formula so obtained is identical with that which results from Sommerfeld's relativistic energy formula\(^{74}\)) in the old quantum mechanics, with the replacement of the azimuthal quantum number $n$ by $n + 1/2$ and the radial quantum number $n'$ by $n' - 1/2$. Fock's energy level formula gives, for the splitting in the fine structure of energy levels, much larger values than the experimental ones.\(^{75}\)) This problem comes to be solved when Dirac's relativistic wave equation of electron is presented (cf. §7 of the present section).

Gordon obtains, in his paper "The Compton Effect according to Schrödinger's Theory"\(^{76}\)) received on Sept. 29, 1926, the same relativistic wave equation in the presence of an electromagnetic field as that by Schrödinger by the same procedure as the one taken by Schrödinger in the fourth paper, to which Gordon refers. Gordon uses the relativistic notations to put $x_4 = ict$ and $\Phi_4 = i\Phi_0$, where $\Phi_0$ is the scalar potential representing the electromagnetic field together with the components of the vector potential $\Phi_1, \Phi_2$ and $\Phi_3$, and writes down the wave equation as

$$\left[ \sum_{\alpha=1}^{4} \left\{ \left( \frac{\hbar}{2\pi i} \right) \frac{\partial}{\partial x_\alpha} - \left( \frac{e}{c} \right) \Phi_\alpha \right\}^2 + m^2 c^2 \right] \psi = 0. \quad (144)$$

Gordon shows further that the quantity

$$j_\alpha = -i \left\{ \bar{\psi} \frac{\partial \psi}{\partial x_\alpha} - \psi \frac{\partial \bar{\psi}}{\partial x_\alpha} - \left( \frac{4\pi ie}{\hbar c} \right) \Phi_\alpha \psi \bar{\psi} \right\}, \quad (145)$$

which is determined in terms of $\psi$ and its complex conjugate $\bar{\psi}$, satisfies the continuity equation $\sum_{\alpha=1}^{4} \partial j_\alpha/\partial x_\alpha = 0$,\(^{77}\)) and interprets $j_k (k = 1, 2, 3)$ as "the electric current density" and $\rho$ defined by $j_4 = ict\rho$ as "the electric charge density". He also shows that the wave equation (144) can be derived from the

---


\(^{75}\)In some present-day textbooks on quantum mechanics, this formula derived by Fock is called Schrödinger's relativistic energy level formula.

\(^{76}\)W. Gordon, Zeit. f. Phys. 40 (1926), 117.

\(^{77}\)He makes use of the condition $\sum_{\alpha=1}^{4} \partial \Phi_\alpha/\partial x_\alpha = 0$, which the electromagnetic potentials must obey.
variation of a quadratic form of $\psi$ and $\bar{\psi}$, in a manner similar to the variational method used in Schrödinger's first paper.

Klein obtains, too, in his paper "Electromagnetics and Wave Mechanics from the Standpoint of the Correspondence Principle"\footnote{O. Klein, Zeit. f. Phys. 41 (1927), 407.} received on Dec. 6, 1926, the same continuity equation as above by referring to Schrödinger's papers up to the fourth and by making use of a procedure similar to Gordon's. Because of the presentation by Gordon and Klein of the continuity equation together with the relativistic wave equation, Eq. (144) for the case of no potentials $\Phi_\alpha$ is nowadays often called the Klein-Gordon equation.

Dirac attempts, in his paper "Relativity Quantum Mechanics with an Application to Compton Scattering"\footnote{P. A. M. Dirac, Proc. Roy. Soc. A111 (1926), 405.} received on Apr. 29, 1926, to formulate relativistically matrix mechanics. As we have seen in Sec. 3, §2, Chap. 1 of the present volume, Dirac formulated matrix mechanics from the standpoint of "$q$-algebra" by making use of "uniformising variables", which consist of pairs of an action variable and its canonically conjugate angle variable, for the case that the Hamiltonian is a function only of the action variables, and showed that the values allowed for an action variable differ from one another by integer multiples of $\hbar$. In the case of the Compton scattering by a free electron, "there is no restriction on the values that the action variable can take" because "the initial value of the action variable is determined by the initial velocity". In consideration of the fact that the Hamiltonian in this case contains the time, Dirac makes a relativistic formulation of matrix mechanics.

Nothing that the time $t$ and the negative of the energy, $-E$, form a pair of canonically conjugate variables in classical mechanics, Dirac assumes the commutation relation $[t, -E] = -i\hbar$. He points out that his commutation relation becomes $[x_4, p_4] = i\hbar$, which is of the same form as those for the space coordinates and momenta, if one puts $x_4 = ict$ and $p_4 = iE/c$. Dirac applies these commutation relations to the Hamiltonian for a charged particle in an electromagnetic field

$$
H = \frac{1}{2m} \left[ \sum_{\alpha=1}^{4} \left\{ p_\alpha - \left( \frac{e}{c} \right) \Phi_\alpha \right\}^2 + m^2 c^2 \right], \quad (146)
$$

to treat the Compton effect.
§2.2 Compton effect

If the incident radiation is a plane wave of frequency $\nu_1$ propagating in the direction of the $x_1$-axis with its electric vector in the direction of the $x_2$-axis, one has $\Phi_2 = a \cos[2\pi \nu_1(ct - x_1)]$, $\Phi_3 = \Phi_4 = 0$, where $a$ is a constant. The intensity of the incident radiation is given by $I_0 = \pi c^2 a^2 \nu^2/2$. Supposing that the radiation is emitted through the Compton effect in a direction defined by the direction cosines $l_1, l_2, l_3$, Dirac applies the canonical transformation in classical theory given by

$$x'_1 = ct - x, \quad x'_2 = x_2, \quad x'_3 = x_3, \quad t' = t - (l_1 x_1 + l_2 x_2 + l_3 x_3); \quad p_1 = -p'_1 + l_1 E'/c, \quad p_2 = p'_2 + l_2 E'/c, \quad p_3 = p'_3 + l_3 E'/c, \quad -E = -E' + cp'_1.$$  

He calculates then the transformed Hamiltonian $H'$ by neglecting $a^2$. In the application of the commutation relations between $x'$ and $p'$, etc., there is the problem of the order of non-commutable variables. Dirac shows, however, that the frequency and intensity of the emitted radiation are not dependent on those orders.

Because this $H'$ does not contain $t'$ explicitly, it is possible to find the action variable and angle variable $w'$ that become "uniformising variables". Also $H'$ does not contain $x'_2$ and $x'_3$, so that one has the selection rules $\Delta p'_2 = 0$, $\Delta p'_3 = 0$ and $\Delta J' = -\hbar$. Combining these with Bohr's frequency relation $\Delta E' = -2\pi \hbar \nu'$, where $\nu'$ is the frequency of the emitted radiation, to calculate $\Delta p_1$ and $\Delta E$ from the above-mentioned canonical transformations, one can derive "the conservation of momentum and energy on Compton's light quantum theory of scattering".

Dirac expresses then the components of the electric dipole moment, which are perpendicular to the direction of the emitted radiation, in terms of $p'_2, p'_3, J$ and $w'$ by arranging the equations derived from $[x'_1, H']$, etc. The coefficients of their terms proportional to $e^{\pm iw'}$ determine the intensity of the emitted radiation. With the use of the initial value of $J'$ corresponding to $p_1 = p_2 = p_3 = 0$ and $E = mc^2$, Dirac obtains an intensity of the scattered radiation that is $(\nu'/\nu)^3$ times the value according to classical theory, that is, the intensity of the Thomson scattering. In consideration of the variation of $\nu'/\nu$ with the angle of scattering, Dirac says that "this result is not very different from Compton's formula", and that "in particular, they agree when the angle of scattering is $0^\circ$ and $180^\circ$".

The integration of the intensity multiplied by $\nu'/\nu$ over all solid angles gives the absorption coefficient $\sigma$. Dirac's result is expressed as

$$\sigma/\sigma_0 = (1 + \alpha) \left( \frac{3}{4\alpha^2} \right) \left\{ \frac{2(1 + \alpha)}{(1 + 2\alpha)} - \frac{1}{\alpha} \log(1 + 2\alpha) \right\}, \quad (147)$$
where $\sigma_0$ is the absorption coefficient in classical theory and $\alpha = h\nu/mc$. Dirac concludes that “for ordinary values of $\alpha$ this lies very close to Compton’s expression” $\sigma/\sigma_0 = 1/(1 + 2\alpha)$. For $\alpha = 1$ Eq. (147) gives “a result 5.7% greater than Compton’s”. For small values of $\alpha$ Eq. (147) give $\sigma/\sigma_0 \approx 1 - 2\alpha + 21\alpha^2/5$, which differs from Compton’s formula in the second order of $\alpha$ with respect to deviation from the classical theory. For large values of $\alpha$ it behaves asymptotically as $\sigma/\sigma_0 \approx 1/\alpha$ differently from Compton’s formula.

Gordon says, in his paper cited above in i), that “the Compton effect has already been calculated by Dirac according to Heisenberg’s method”, and “now the same problem should be treated after Schrödinger”. He states also that “Schrödinger’s procedure has the merit that the customary mathematical aids serve themselves”. Gordon treats thus the Compton effect on the basis of wave mechanics, and calculates the intensity of the emitted radiation in a correspondence-theoretical way, getting the same result as Dirac. This shows the equality of matrix mechanics formulated relativistically in the above-mentioned way with wave mechanics. However, this shows at the same time that both the theories include problems still to be solved. As we shall see later in §7 of the present section, Klein and Nishina derive a formula of Compton scattering that is different from Dirac’s and Gordon’s, by applying the relativistic wave equation of electron shortly after its proposal by Dirac.

Now, Gordon calculates first the current density $j_k$ with the use of the electromagnetic potentials $\Phi_\alpha$ of an incident plane wave. In this calculation, he determines $\partial \psi/\partial x_k$ appearing in the expression for $j_k$ from the term $\Phi_\alpha \partial \psi/\partial x_k$ in the wave equation, by making the Fourier development of the wave function $\psi$ with respect to the components of the momentum of the electron, $p_1, p_2, p_3$. If the center-of-mass coordinates are denoted by $X_k$, one has the relation $edX_k/dt = \int j_k dx_1 dx_2 dx_3$. Integrating this equation with the use of the $j_k$ so determined and equating the terms in the result which are independent of $\Phi_\alpha$ to the velocity of the electron in the classical theory, $dx_k/dt = c^2 p_k/E$, multiplied by $e$, Gordon determines then the coefficients of the Fourier development. He calculates finally the electromagnetic potentials $\Phi_\alpha$ of the outgoing radiation, in the form of retarded potential emerging from the $j_k$ thus obtained as the sources, to get the frequency and intensity of the radiation emitted into various directions.

---

81) Compton’s formula gives $\sigma/\sigma_0 \approx 1 - 2\alpha + 4\alpha^2$ for small values of $\alpha$, and $\sigma/\sigma_0 = 1/2\alpha$ for large values of $\alpha$. 
§3. Unification of light-quantum theory and electromagnetic wave theory

In the applications of matrix mechanics and wave mechanics to the Compton effect, which we have described above in §1 and §2, the electromagnetic field is treated as an external field. During the time nearly coincident with the time of these applications, the unification of matrix mechanics and wave mechanics was achieved by Born, London, Dirac and Jordan, and quantum mechanics was established, as we have seen in the last section. Schrödinger’s wave function in the multi-dimensional configuration space got rid of its substantialistic limitation to become the essentialistic concept that represents states of every quantum system.

Dirac applies quantum mechanics, in his paper “The Quantum Theory of the Emission and Absorption of Radiation” received on Feb. 2, 1927, to a system of radiation, citing the above-mentioned papers of London, Dirac and Jordan. As a result, he shows that “there is a complete harmony between the wave and light-quantum descriptions of the interaction” of the atom and radiation. That is, the solution is thus achieved at last of the conflict between the wave-nature and particle-nature of the radiation, which was one of the original problems in quantum theory.

Dirac considers in the first place that the problem concerning the light-quantum nature of free radiation is already solved. Namely, he makes a Fourier series development of the radiation in a cavity, and shows that, in terms of the energy $E_r$ and phase $\theta_r$ of the rth component, the Hamiltonian takes the form $H = \sum_r E_r$, and so $E_r$ and $\theta_r$ satisfy the classical canonical equations $\dot{E}_r = -\frac{\partial H}{\partial \theta_r} = 0$ and $\dot{\theta}_r = \frac{\partial H}{\partial E_r} = 1$. Dirac says, citing Born-Jordan’s paper and Born-Heisenberg-Jordan’s paper which we have mentioned in Sec. 2, §6, Chap. 1 of the present volume, that “the assumption of the standard quantum conditions $\theta_r E_r - E_r \dot{\theta}_r = \hbar$, etc., immediately gives light-quantum properties to the radiation”.

As we have described in that section, Born-Heisenberg-Jordan expressed the oscillation of a string as an aggregate of its harmonic eigen-oscillations, and applied matrix mechanics to each harmonic oscillation to calculate the fluctuation in the energy of the string. By analogy to its result, they deduced Einstein’s formula for the fluctuation in energy of the cavity radiation. We have there described this as “the beginning of the quantum theory of field”.

---

Also, in Sec. 4, §1, Chap. 3, Vol. II of the present book, we have mentioned Rubinowicz’s paper in 1917 in which the quantum condition in the old quantum mechanics was applied to the eigen-oscillations of radiation in a cavity with the result $E_r = n_r h\nu_r$. We have assessed it as “the germ of the quantization of field”. Furthermore, in Sec. 1, §3, Chap. 5 of the same volume, we have treated Compton’s work in 1923 in which he applied the quantum condition to a light beam because of its periodicity at every interval equal to the wavelength, deriving the momentum of a light quantum. We have assessed this as “the germ of the quantization of wave field with the periodic boundary condition”.

§3.1 Bose-Einstein statistics and the second quantization

To complete the quantum theory of field, it is necessary to treat the interaction of radiation and an atom, so as to “build up a fairly satisfactory theory of the emission of radiation and of the reaction of the radiation field on the emitting system”. In preparation for the treatment of this problem, Dirac develops the quantum mechanics of “an assembly of particles satisfying the Einstein-Bose statistics”. He writes down “the wave equation” of a particle as

$$\hbar \frac{\partial \psi}{\partial t} = (H_0 + V)\psi, \quad (148)$$

where $H_0$ is the Hamiltonian of the unperturbed system and $V$ denotes some perturbing energy. In terms of the eigenfunctions $\varphi_r$‘s of $H_0$, each satisfying $H_0 \varphi_r = W_r \varphi_r$, he expands $\psi$ as $\sum_r b_r \varphi_r$. If the $b_r$‘s are normalized so that $\sum_r |b_r|^2 = 1$, $|b_r|^2 = b_r b_r^*$ is “the probability of the system being in the state $r$”. Denoting by $V_{rs} = V_{sr}^*$ the usual matrix element of $V$ made with $\varphi_r$ and $\varphi_s$, one sees from Eq. (148) that $b_r$ and $\dot{b}_r$ satisfy the equations

$$i\hbar \dot{b}_r = W_r b_r + \sum_s V_{rs} b_s, \quad -i\hbar \dot{b}_r^* = W_r b_r^* + \sum_s b_s^* V_{sr}, \quad (149)$$

Dirac notes that these equations are identical with the canonical equations (in classical theory) $\dot{b}_r = \partial F_1 / \partial (i\hbar b_r^*)$ and $i\hbar \dot{b}_r^* = -\partial F_1 / \partial b$ for the pair of canonical
variables $b_r$ and $i\hbar \dot{b}_r^*$ with the Hamiltonian\textsuperscript{83)

$$F_1 = \sum_r W_r b_r^* b_r + \sum_{rs} V_{rs} b_r^* b_s.$$  \hfill (150)

If the normalization $\sum_r |b_r|^2 = N$ is chosen, “the theory will apply directly to an assembly of $N$ similar independent systems”.\textsuperscript{84)} In that case, $N_r = |b_r|^2$ is “the probable number of systems in the state $r$. Then, transforming the pair of canonical variables $b_r$ and $i\hbar \dot{b}_r^*$ into a new pair of $N_r$ and $\theta_r$ defined by

$$b_r = N^{1/2} \exp(-i\theta_r/\hbar), \quad b_r^* = N^{1/2} \exp(i\theta_r/\hbar),$$  \hfill (151)

one sees that $N_r$ and $\theta_r$ satisfy the classical canonical equations $\dot{N}_r = -\partial F/\partial \theta_r$ and $\dot{\theta}_r = \partial F/\partial N_r$, with the transformed Hamiltonian

$$F = \sum_r W_r N_r + \sum_{rs} V_{rs} N_r^{1/2} N_s^{1/2} \exp[i(\theta_r - \theta_s)/\hbar].$$  \hfill (152)

Imposing the quantum mechanical commutation relations $b_r (i\hbar b_s^*) - (i\hbar b_s^*) b_r = i\hbar \delta_{rs}$, etc. on $b_r$ and $i\hbar \dot{b}_r^*$, that is, putting

$$b_r b_s^* - b_s^* b_r = \delta_{rs}, \quad b_r b_s - b_s b_r = b_r^* b_s^* - b_s^* b_r^* = 0,$$  \hfill (153)

one sees that $N_r$ and $\theta_r$, which are defined by

$$\begin{aligned}
&\begin{cases}
  b_r = (N_r + 1)^{1/2} \exp(-i\theta_r/\hbar) = \exp(-i\theta_r/\hbar) N_r^{1/2} \\
  b_r^* = \exp(i\theta_r/\hbar) = \exp(i\theta_r/\hbar) (N_r + 1)^{1/2}
\end{cases},
\end{aligned}$$  \hfill (154)

make a pair of canonical variables satisfying the commutation relations, and also that $N_r$ has only the numbers $0, 1, 2, \ldots$ for its eigenvalue, as in the case

\textsuperscript{83)Schrodinger presented a kind of thinking similar to Dirac's in the paper “The Energy-Momentum Law of Matter Wave” [Ann. d. Phys. 82 (1927), 265], received about two months before Dirac's paper under consideration, saying that Gordon's method of deriving the relativistic wave equation with the use of the variation principle, which we have described in §2.1 of the present section, was "analogous to the method of representation in the mechanics of mass points with the use of Lagrangian". Schrödinger gave then the Lagrangian of the system composite of "matter field" and "electromagnetic field", from which he derived classical-theoretically some relations. Confer also Sec. 1, §3, Chap. 2 of the present volume.

\textsuperscript{84)In Dirac's paper about five months earlier, which we have treated in Sec. 1, §6 of the present chapter, this normalization was only taken, and the transition probability of induced emission per individual atom was calculated as an average value over many atoms. Dirac considers now that Eq. (148) is for a single particle, in accordance with the establishment of quantum mechanics."
of the quantum number of a harmonic oscillator. $F$ given by Eq. (152) is then transformed into

$$ F = \sum_r W_r N_r + \sum_{rs} V_{rs} N_r^{1/2} (N_s + 1 - \delta_{rs})^{1/2} \exp[i(\theta_r - \theta_s)/\hbar]. \quad (155) $$

"The wave equation written in terms of the variables $N_r$" becomes

$$ i\hbar (\partial / \partial t) \Psi(N_1, N_2, N_3, \ldots) = F \Psi(N_1, N_2, N_3, \ldots), \quad (156) $$

where $\theta_r$ occurring in $F$ is set to be $\theta_r = -i\hbar \partial / \partial N_r$. With the normalization $\sum_{N_1, N_2, \ldots} |\Psi(N_1, N_2, \ldots)|^2 = 1$, $|\Psi(N_1, N_2, \ldots)|^2$ is the probability of the "distribution" in which $N_1$ systems are in state 1, $N_2$ systems in state 2, ... Because $\exp(\pm i\hbar \partial / \partial N_r) f(N_r) = f(N_r \pm 1)$ for a function $f(N_r)$ of $N_r$, $F$ is the sum of the terms increasing $N_r$ by unity and decreasing $N_r$ by unity.

In this way, Dirac arrived at "the wave equation" for "an assembly of particles satisfying the Einstein-Bose statistics". Regarding the Schrödinger equation for a one-body problem, i.e. Eq. (148), as the classical equation for the de Broglie wave, Dirac applied to it the result of quantum mechanics which connects Schrödinger's wave function in a multi-dimensional configuration space, grasped as the transformation function, with the probability. According to v. Neumann's paper, the coefficients $b_r$'s in the development of $\psi$ are the components of the state vector in the abstract Hilbert space for a particle, and $\Psi(N_1, N_2, \ldots)$'s with the values of $N_1, N_2, \ldots$, each being given any one of $0, 1, 2, \ldots$, are the components of the state vector in the abstract Hilbert space for a system of particles obeying the Bose-Einstein statistics.

The Planck constant $\hbar$ included in Eq. (148) represents the wave nature, that is the quantum-mechanical attribute, of a particle. Dirac's procedure of applying the commutation relations involving $\hbar$ to Eq. (148), which is regarded as a classical wave equation, is later known as the second quantization. Equation (156) also becomes known as the equation of field in the number representation.

Now, in order to treat the interaction of "an assembly of identical systems satisfying the Einstein-Bose statistics" with an atom, Dirac adds to $F$ given by Eq. (155) the Hamiltonian of the atom $H_P(J)$, where $J$ denotes collectively the action variables determining the state of the atom. $V_{rs}$ is thus considered to be a matrix element with respect to $J$ and is written as $V_{rs}(J, J')$. $\Psi$ in Eq. (156) is accordingly expressed as $\Psi(J, N_1, N_2, \ldots)$. Dirac applies Eq. (156) so reformed to the interaction of the radiation with the atom, and considers it
as the description of the interaction “from the light-quantum point of view”. Namely, he regards the radiation as “an assembly of particles moving with the velocity of light and satisfying the Einstein-Bose statistics”.

§3.2 Application to the system of light quanta

Dirac says that “the light-quantum has the peculiarity that it apparently ceases to exist when it is in one of its stationary states, namely, the zero state, in which its momentum, and therefore also its energy, are zero”. He then says that “when a light quantum is absorbed it can be considered to jump into this zero states, and when one is emitted it can be considered to jump from the zero state to one in which it is physically in evidence, so that it appears to have been created”. Because Dirac bases his study on Schrödinger’s wave equation, he understands that light quantum is a kind of permanent existence like a material particle such as the electron.

“Since there is no limit to the number of light quanta that may be created in this way, we must suppose that there are infinite number of light quanta in zero state”. So, Dirac considers the vacuum like the state of Bose-Einstein condensation at the absolute zero-temperature in Einstein's theory of quantum gas, which we have treated in Sec. 2, §4, Chap. 5, Vol. II of the present book.

Because $N_0$ in the Hamiltonian $F$ given by Eq. (155) is now taken to be infinity, $V_{0r}$ and $V_{r0}$ must be infinitesimal in comparison with it. Dirac thus sets

$$V_{0r}(N_0 + 1)^{1/2} \exp(-i\theta_0/\hbar) = V_r, \quad V_{r0}N_0^{1/2} \exp(i\theta_0/\hbar) = V_r^*, \quad (157)$$

supposing that $V_{0r}N_0^{1/2}$ and $V_{r0}N_0^{1/2}$ are finite. Because $N_0$ and $\theta_0$ are “practically constants”, $V_r$ and $V_r^*$ depend only on the canonical variables of the atom.85)

On the other hand, $V_r$ and $V_{rs} (r \neq s)$ are “unknown interaction coefficients” in the description “from the light quantum point of view”, so that they are not concretely determined. Dirac says, however, that “the wave point of view gives values for them”. Namely, he resolves the radiation into its Fourier components. Each component $r$ can be described by a vector potential $\kappa_r$ chosen so as to make the scalar potential zero. According to classical theory, the interaction energy between the radiation and the atom is given by $\sum_r (\kappa_r/c) \dot{X}_r$,

85) Neglecting quantities of the order of $N_0^{-1/2}$, one sees that $\dot{\theta}_0 = \partial F/\partial N_0 \sim W_0 = 0$ and therefore $\theta_0$ is practically constant.
where $X_r$ the component of the total polarization of the atom in the direction of $\kappa_r$, which is the direction of the electric vector of the component $r$.

As we have seen at the beginning of the present subsection §3, the energy $E_r$ and phase $\theta_r$ of the component $r$ constitute a pair of canonical variables. Dirac takes $N_r = E_r/(2\pi \hbar \nu_r)$ as the canonical variable substituting for $E_r$ and puts $\kappa_r = a_r \cos(\theta_r/\hbar)$, denoting the amplitude of $\kappa_r$ by $a_r$. The flow of energy per unit area per unit time gives then the relation $a_r = 2(\hbar \nu_r/c \sigma_r)^{1/2} N_r^{1/2}$, where $\sigma_r$ is the number of the freedom of polarization. “On the quantum theory we must make the variables $N_r$ and $\theta_r$ canonical $q$-numbers like the variables $J_k$, $w_k$ that describe the atom”. Dirac makes thus $N_r^{1/2} \cos(\theta_r/\hbar)$ symmetrical to put $\{N_r^{1/2} \cos(i\theta_r/\hbar) + \exp(-i\theta_r/\hbar)N_r^{1/2}\}$, and puts $\exp(-i\theta_r/\hbar)N_r^{1/2} = (N_r + 1)^{1/2} \exp(-i\theta_r/\hbar)$ as in Eq. (154). He gets then from $\sum_r (\kappa_r/e \dot{X}_r)$

$$V_r = V_r^* = (\hbar \nu_r/c^3 \sigma_r)^{1/2} \dot{X}_r, \quad V_{rs} = 0 \quad (r, s \neq 0), \quad (158)$$

in terms of which $F$ is given to be

$$F = H_p(J) + \sum_r (2\pi \hbar \nu_r) N_r + \sum_r \times \{V_r N_r^{1/2} \exp(i\theta_r/\hbar) + V_r^*(N_r + 1)^{1/2} \exp(-i\theta_r/\hbar)\}. \quad (159)$$

Thus, Dirac concludes that “the wave point of view is consistent with the light quantum point of view”, by showing that “the Hamiltonian which describes the interaction of the atom and the electromagnetic waves can be made identical with the Hamiltonian for the problem of the interaction of the atom with an assembly of particles moving with the velocity of light and satisfying the Einstein-Bose statistics”.

It should be noted that Dirac treats the problem of the interaction of atom and radiation by assuming that the problem of the light-quantum nature of free radiation field has been already solved. This is because Einstein’s spontaneous transition probability is shown to be derived from the fundamental equation, without any additional assumption. Namely, Dirac calculates the time dependence of $\Psi(J, N_1, N_2, \ldots)$ with the use of $F$ given by Eq. (159), by applying a method which is equivalent to the first approximation in the time-dependent perturbation, so called in the present-day quantum mechanics.\textsuperscript{86} He derives

\textsuperscript{86}His method of calculation is substantially identical with Born’s method of approximation which we have described in §1 of the last section. As we have mentioned there, however, Born treated the radiation as “an irregular sequence of many wave pulses”, not arriving at the quantization of radiation field.
the transition probability not only for induced emission and absorption but also for spontaneous emission, giving "a proper account of spontaneous emission".

Dirac says also that "the theory enables one to understand how it comes about that there is no violation of the law of the conservation of energy when, say, a photo-electron is emitted from an atom under the action of extremely weak incident radiation". This is to point out the fact that the energy of a photo-electron is independent of the intensity of the incident light, being given by the formula: $h\nu - \text{work function}$.

As we have described in Sec. 3, §3, Chap. 1, Vol. II of the present book, Einstein pointed out in 1909 concerning the production of cathode rays by X-rays that "secondary cathode rays of velocity of the same order of magnitudes as that of the primary cathode rays" are emitted by X-rays which "seem to be directional", even though the intensity of X-rays is diminished by weakening that of the primary cathode rays, and emphasized the light quantum theory. On the other hand, Lorentz discussed in 1910 the absorption of light, saying that one should imagine that a molecule receives temporally the energy to catch it if the molecule gathers a whole quantum, or releases it if the molecule does not succeed in doing so, and criticized the light-quantum theory. And, as we have seen in Sec. 3, §§1 and 2, Chap. 5 of the same volume, Bohr in 1923, citing this paper of Lorentz, doubted the conservation of energy and momentum in the interaction between radiation and an atom, and proposed in 1925 with Kramers and Slater the statistical conservation of energy and momentum, which was immediately negated by Compton-Simon's experiment.

Now, according to Dirac, "the energy of interaction of the atom and the radiation is a q-number that does not commute with the first integrals of the motion of the atom alone or with the intensity of the radiation". Thus, "one cannot specify this energy by a c-number at the same time that one specifies the stationary state of the atom and the intensity of the radiation by c-numbers". Dirac states that "in particular, one cannot say that the interaction energy tends to zero as the intensity of the incident radiation tends to zero", and "there is thus always an unspecifiable amount of interaction energy which can supply the energy for the photo-electron". This is to point out the quantum-mechanical characteristic of the interaction of the atom and the radiation. On account of this characteristics, discussions based on the classical theory such as Lorentz's are found not to hold. This characteristic is connected with the stereo-structure of the logic of quantum mechanics, which we shall describe in the next chapter.
§4. Introduction of spin wave functions

§4.1 Darwin's de Broglie wave of vector type

As we have described in Sec. 2, §5, Chap. 1 of the present volume, Heisenberg and Jordan applied matrix mechanics to Uhlenbeck and Goudsmit's model of spinning electron to treat the anomalous Zeeman effect of hydrogen-like atoms, and derived Landé's $g$-factor by making use of Born-Heisenberg-Jordan's result in matrix mechanics for the angular momentum in general and assuming the anomalous coefficient 2 for the magnetic dipole moment of the spin angular momentum.

Darwin tries in his paper "The Electron as a Vector Wave",\(^\text{87}\) issued on Feb. 19, 1927, to construct "equations of the wave form to conform to classical dynamics in the limit", in order to bring the spin into "wave mechanics". He says that a way out of the difficulty that "the wave mechanics definitely excludes half quantum numbers for the spin, and so would lead to triplets instead of doublets — 1, 0, −1 instead of $1/2, −1/2$", is to "assimilate the electron to a transverse rather than a longitudinal wave, for this at once provides the number of states with the necessary factor 2". In the case of an electromagnetic wave, the equations $\text{div} \ E = \text{div} \ B = 0$ ensure that it is a transverse wave. Since "the Schrödinger function will not tolerate a divergence relation of any kind", Darwin goes "to construct by a much more inductive process a system of waves, of a vector character though not transverse". Namely, taking the relativistic wave equation $Df = 0$ of the Klein-Gordon type as "Schrödinger's equation", he assumes that the equations

\begin{equation}
Df + \alpha f + \beta g = 0, \quad Dg + \gamma f + \delta g = 0
\end{equation}

(160)

hold simultaneously for $f$ and another wave function $g$.\(^\text{88}\) In these equations, $\alpha, \beta, \gamma, \delta$ are "small perturbing operators which will give the observed values for doublet spectra".

Taking the azimuthal quantum numbers $k = 0, 1, 2, \ldots$, for the $s, p, d, \ldots$ terms respectively, Darwin finds by trial $\alpha, \beta, \gamma, \delta$ which reproduce Landé's formula of spectra. Transforming Eq. (160) into equations for functions


\(^{88}\)Darwin cites Dirac's paper mentioned in §2 of the present section and includes into $D$ the terms of interactions of the electron with the Coulomb field and the external magnetic field. In the process of calculation, he applies the nonrelativistic approximation $i(h/2\pi)\partial/\partial t \simeq mc^2$. 

The Establishment of Quantum Mechanics

$X_1, X_2, X_3, X_4$, with the use of a process which “is nearly, but not quite, the same as taking real and imaginary parts” of Eq. (160), he shows that the transformed equations “are now in vector form for space transformations, regarding $X_4$ as a scalar and $X_1, X_2, X_3$ as a vector”. As Darwin says, however, “the idea involves difficulty when more than one electron is present”. In this sense Darwin’s wave functions are for de Broglie wave of vector type.

§4.2 Pauli’s wave functions for spin

About two and a half months later than Darwin’s, introduction of two wave functions $f$ and $g$ because of the number of freedom 2 of the electron spin, Pauli introduces the freedom of the spin into the wave function with a method different from Darwin’s in a paper “On the Quantum Mechanics of Magnetic Electron”\textsuperscript{89} received on May 3, 1927. On Darwin’s attempt Pauli says “There arise however difficulties which are connected just again with the number 2 of the positions of the electron in an external field, and of which I do not believe that they can be overcome”.

In classical theory, the mechanical behavior of the angular momentum of an electron is described by the following two pairs of canonical variables. One is that of the magnitude $s$ of the eigen angular momentum $s$ of the electron and the angle $\chi$ of rotation around its axis. The other is that of the component $s_z$ in the direction of the fixed $z$-axis and the azimuthal angle $\varphi$ of $s$. Because $\chi$ is periodic and does not appear in the Hamiltonian, $s$ is a constant and one is left with only the pair of canonical variables $(s_z, \varphi)$.

Pauli says that “with the application of Schrödinger’s original method, one would have an eigen function, which depends on the angle $\varphi$ aside from the three position coordinates of the center of mass of the electron (denoted shortly by $q_k$ or also $q$)”.

Then it gives

$$|\psi_E(q, \varphi)|^2 dq_1 dq_2 dq_3 d\varphi$$

for the probability that, in the quantum state concerned of energy $E$, the position coordinates lie in the intervals $q_k, q_k + dq_k$ and the angle $\varphi$ lies in $(\varphi, \varphi + d\varphi)$. Accordingly “$s_z$ would be applied by means of $(\hbar/2\pi i)\partial/\partial \varphi$ to the eigenfunction $\psi$”.

Pauli points out however that “the circumstance that the number of the quantum-theoretically allowed orientations of the angular momentum of the

\textsuperscript{89} W. Pauli, Zeit. f. Phys. 43 (1927), 601.
Developments in Various Problems

215

electron is 2, leads to the consequence that the function \( \psi_E(q, \varphi) \) so defined does not come back to its initial value, but changes its sign with the continuous change of \( \varphi \) from 0 to \( 2\pi \). Thus, Pauli "avoids the occurrence of such double-ness, by introducing as the independent variable the component \( s_z \) of the angular momentum instead of \( \varphi \)." In quantum mechanics \( s_z \) can take "the characteristic values" \( \pm (h/2\pi)/2 \). "This should mean that the function \( \psi_E(q, s_z) \) decomposes into the two functions

\[
\psi_{\alpha, E}(q_k) \quad \text{and} \quad \psi_{\beta, E}(q_k),
\]

which correspond to the values \( s_z = (h/2\pi)/2 \) and \( s_z = -(h/2\pi)/2 \).

\[ |\psi_{\alpha, E}(q_k)|^2 dq_1 dq_2 dq_3 \text{ gives "the probability that in the stationary state concerned } s_z \text{ has the value } (h/2\pi)/2 \text{ at the same time as } q_k \text{ lie in } (q_k, q_k + dq_k). \]

\[ |\psi_{\beta, E}(q_k)|^2 dq_1 dq_2 dq_3 \text{ gives a similar probability for } s_z = -(h/2\pi)/2. \]

Regarding the variable \( \varphi \), Pauli mentions that one follows the procedure to operate on \( \psi(q_k, s_z) \) the Hamiltonian, which is expressed in terms of the pairs of canonical variables \((p_k, q_k)\) and \((s_z, \varphi)\) with \( p_k \) and \( \varphi \) replaced by \( (h/2\pi)i \partial/\partial q_k \) and \((h/2\pi)i \partial/\partial s \) respectively, and "finally to pass over the limit where \( \psi \) is different from zero only for \( s_z = (h/2\pi)/2 \) and \( s_z = -(h/2\pi)/2 \)."

Because the components of the spin angular momentum satisfy the commutation relations

\[
s_x s_y - s_y s_x = 2 i s_z, \ldots; \quad s_x^2 + s_y^2 + s_z^2 = 1,
\]

when they are measured in unit of \( (h/2\pi)/2 \), one has the equations

\[
\begin{align*}
\begin{cases}
 s_x \psi_\alpha = \psi_\beta, & s_y \psi_\alpha = -i \psi_\beta, & s_z \psi_\alpha = \psi_\alpha, \\
 s_x \psi_\beta = \psi_\alpha, & s_y \psi_\beta = i \psi_\alpha, & s_z \psi_\beta = -\psi_\beta.
\end{cases}
\end{align*}
\]

Pauli points out that these can be expressed "in the symbolic matrix form"

\[
\begin{align*}
 s_x \psi &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \psi, &
 s_y \psi &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \psi, &
 s_z \psi &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \psi,
\end{align*}
\]

where \( \psi \) is considered to consist of two components \( \psi_\alpha \) and \( \psi_\beta \). These matrices later become known as the Pauli spin matrices.

\[ ^{90} \text{Though Pauli does not say in addition about } s_z, \text{ Schrödinger's method gives the equation } s_z \psi(\varphi) = \pm \left( (h/2\pi)/2 \right) \psi(\varphi), \text{ whose solutions } \psi(\varphi) = \text{const.} \exp(\pm i\varphi/2) \text{ result in the relation } \psi(2\pi) = -\psi(0). \]
On the transformation of the spin matrices accompanying the rotation transformation of the system of coordinates, he makes analyses similar to those seen in present-day textbooks on quantum mechanics. He then treats the anomalous Zeeman effect of a hydrogen-like atom. Taking for the perturbation Hamiltonian “the interaction term of the eigen magnetic moment with the external magnetic field”, and “the interaction term of a moving electron having a magnetic moment with the Coulomb electric field”, that is what one nowadays calls spin-orbit interaction, Pauli uses the nonrelativistic Schrödinger equation for \( \psi \) to derive simultaneous differential equations for \( \psi_\alpha \) and \( \psi_\beta \). This set of equations is identical with what is also seen in present-day textbooks. Pauli says that the solution to the equations “leads to no new thing in comparison with the results of Heisenberg and Jordan”.

Pauli says that, in the case of the external magnetic field being parallel to the \( z \)-axis, the simultaneous equations for \( \psi_\alpha \) and \( \psi_\beta \) “go over to the one that has already been set up by Darwin”. Pauli stresses however that “in contrast to Darwin” he sees “as the source of these equations the commutation relations, not the idea that the amplitudes of the de Broglie wave are directional quantities”.

Pauli’s theory can be applied to the case of many electrons, though Darwin’s attempt contained difficulties in that case. If there are two electrons, for example, it suffices to consider the set of four functions \( \psi_{\alpha_1 \alpha_2}(q_1, q_2), \psi_{\alpha_1 \beta_2}(q_1, q_2), \psi_{\beta_1 \alpha_2}(q_1, q_2), \psi_{\beta_1 \beta_2}(q_1, q_2) \) where the suffix \( \alpha_1, \beta_2 \) indicates that electron 1 is in the state \( \alpha \) and electron 2 in the state \( \beta \), and so on. Furthermore the eigenfunctions can be made, “according to Heisenberg and Dirac”, to satisfy “the symmetry property for exchange of two electrons”.

Pauli’s idea of introducing the spin wave functions brings results in systematizing the formulation of quantum mechanics, but it still contains some difficulties. Namely, as Pauli says, “the equations are not written relativistically invariant and, in the case of the hydrogen atom, are valid only in the approximation in which the dynamical behavior of the eigen moment can be considered on the average over the orbital motion”.

This means that use is made of the nonrelativistic Schrödinger equation and, as in Heisenberg-Jordan’s paper, of the substantialistic spin-orbit interaction which was derived by Thomas-Frenkel from quantum condition for nonrelativistic orbital motions together with the relativistic kinematics.\(^{91}\) Pauli’s

\(^{91}\) Cf. Sec. 3, §3, Chap. 4, Vol. II of the present book.
study of the transformation of the spin accompanying the transformation of coordinates is also restricted to the case of the rotation of coordinates, and is not applied to the relativistic Lorentz transformation. These difficulties in Pauli's theory, which he himself considers "provisional and approximate", are solved by the Dirac relativistic wave equation, which we shall treat in §7 of the present section.

Pauli's study uses, for the angle variable $\varphi$ canonically conjugate to $s_z$, the process of taking the limit mentioned above after replacement of $\varphi$ with $(\hbar/2\pi i)\partial/\partial s_z$. Namely, $s_z$ and $\varphi$ are supposed to obey the ordinary commutation relation $s_z \varphi - \varphi s_z = i\hbar$. Concerning this problem Jordan says, in his paper "On a New Establishment of Quantum Mechanics II"\(^{92}\) received on Jun. 3, 1927, that "we must conclude that in the case of the magnetic electron anything analogous to such an angle of rotation does not exist". As a reason for this, Jordan states that because $s_z$ takes only two values, "the necessary orthogonality of the amplitude belonging to $s_z$ and its hypothetical conjugate can be established only when this conjugate quantity possesses also merely two (not continuously many) eigenvalues".

"Wishing to make really as the starting-point of our considerations the following assumption", Jordan supposes that "in the case of the magnetic electron one cannot measure at all quantities other than the component $s_z$ in any direction $z$ of the angular momentum". This should be considered as emphasizing that the electron spin, unlike a kinematical quantity such as orbital angular momentum, is a quantum mechanical belonging of the electron. This problem is also solved by the Dirac relativistic wave equation, which includes the freedom of the spin.

§5. Symmetries of wave function and atomic and molecular structures

§5.1 Symmetries of wave function and spectral multiplets

Saying that "according to Heisenberg and Dirac" the eigenfunctions can be made to satisfy the symmetry property, as has been described above, Pauli cites also Heisenberg's paper "Many-Body Problems and Resonance in the Quantum Mechanics II"\(^{93}\) received on Dec. 22, 1926. This second paper is

a sequel to his paper written half a year before, which we have mentioned in Sec. 1, §§1 and 2 of the present chapter. In the second paper, Heisenberg cites papers worked out separately by London, Dirac and Jordan later than his first paper, which have been treated in Sec. 2 of the present chapter, to say that such a matrix $S$ [with elements denoted by $S_n(q)$]$^94$ of canonical transformation, which has a discrete index of row and a continuous index of column, gives “Schrödinger’s eigenfunctions”. Heisenberg cites Dirac’s and Jordan’s papers mentioned above as yet “to appear”.

As we have described before, in this paper of Jordan, Pauli’s paper, which was received four months later than Jordan’s, was cited as “in print”, and it was stated that Pauli generalized the wave function $\varphi_n(q)$ to introduce the probability amplitude $\varphi(q, \beta_0)$ in the case that a dynamical variable $\beta$ other than $q$ takes a value $\beta_0$. Pauli used the component $s_z$ of the spin for $\beta$.

Heisenberg considers, in the second paper under consideration, that the matrix $S$ of canonical transformation becomes, in the treatment of the interaction of the magnetic moment and center-of-mass motion of the electron, the product of $S(\xi_1, \xi_2, \xi_3)$, which depends on the center-of-mass coordinates $(\xi_1, \xi_2, \xi_3)$, and $S^4(\xi_4)$, which depends on the coordinate $\xi_4$ describing “the eigen rotation”. He says that if the rotation angle of the spin is chosen as $\xi_4$ the Schrödinger theory would be required to determine $S^4$, but “this problem has still certain difficulties”. As we have mentioned before, Pauli pointed out the difficulty concerning the occurrence of the double-ness of the wave function in this case. Saying that the choice of $\xi_4$ is “completely free” and the quantum mechanical formulation of the spinning electron in words of matrices is “well known”,$^95$ Heisenberg takes $s_z$ for $\xi_4$. $S^4(\xi_4)$ is then given in terms of the two eigenfunctions corresponding to the two values taken by $s_z$. Heisenberg denotes these eigenfunctions by $\alpha$ and $\beta$.

With the eigenfunctions of an electron including its spin in this way, Heisenberg discusses spectral terms of a system of many electrons. If the interaction among the electrons is treated as a perturbation, the function

$$u = l_1 m_2 \cdots p_n$$

(163)

is an unperturbed eigenfunction of the system, where $l_1 m_2 \cdots p_n$ are the eigenfunctions of the electrons 1, 2, \ldots, $n$ respectively. Those functions obtained

$^94$ $n$ and $q$ stand for the quantum numbers of a stationary state and position coordinates respectively.

$^95$ In matrix mechanics, $S^2$ and $s_z$ can be simultaneously diagonalized.
Developments in Various Problems

from $u$ by permuting the numbers $1, 2, \ldots, n$ of the electrons are the unperturbed eigenfunctions as well, and are degenerate in energy. In total, there are $n!$ degenerate eigenfunctions for the unperturbed system. Making use of the properties of the permutation group, Heisenberg classifies the linear combinations of these degenerate eigenfunctions into various systems of spectral terms, showing that different systems of spectral terms are not combinative with each other.$^{96}$

"The eigenfunction of the total system should be antisymmetric with respect to the electrons". So, such linear combinations of the eigenfunctions belonging to a spectral term that satisfy the anti-symmetric property for permutations of two electrons must be looked for. Heisenberg gives concrete expressions of the linear combinations for a three-electron system and treats "the quantum-mechanical resonance effect", that is, the splitting of the degeneracy on account of a perturbation, getting the energies of each spectral term as functions of the matrix elements $W_{\text{imp},\text{imp}}, W_{\text{imp},m\text{p}}, W_{\text{imp},p\text{m}}, W_{\text{imp},l\text{pm}}$ of the interaction energy. Results equivalent to Heisenberg's are also obtained independently by Wigner.$^{97}$

§5.2 Symmetric properties of the wave function of diatomic molecule
and the nuclear spin

Heisenberg discusses then the relation of the symmetric property of eigenfunctions to the structure of band spectra for diatomic molecules with identical nuclei, that is, diatomic molecules consisting of identical atoms. The wave function of a diatomic molecule is the product of the wave functions representing the motion of the electrons, the relative motion of the two nuclei, the rotational motion of the molecular axis, and the center-of-mass motion of the molecule. Now, in the case of a diatomic molecule with identical nuclei, the wave function of the electrons should be either symmetric or antisymmetric with respect to the interchange of the nuclear coordinates. Heisenberg "adopts" this conclusion from Hund's paper "to be published", and thanks Hund for "many discussions on this problem".

96) If $\gamma$ and $\gamma'$ denote two such linear combinations belonging to systems different with each other, and $f$ is a symmetric function of the electrons, it can be shown that the matrix element $\int \gamma^* \gamma' f d\xi$ is zero.

97) E. Wigner, Zeit. f. Phys. 40 (1926), 492. [In this paper of Wigner no consideration is given to spin.]
Hund's paper cited here is "On Interpretation of Molecular Spectra I", received on Nov. 19, 1926, a month earlier than Heisenberg's under consideration. In this paper, Hund discusses the behavior of the wave function of the electrons "qualitatively" on the basis of "the adiabatic theorem". Namely, he arrives at the conclusion mentioned above by considering an adiabatic change of the internuclear distance and noting that the Schrödinger equation for the electrons is invariant for the interchange of the nuclei, and that the system is nearly represented by two isolated atoms when the internuclear distance is large enough and by a single atom with a nucleus having the electric charge twice as large as that of the molecular atom when the internuclear distance is short enough.

Next, Heisenberg says in his paper that the eigenfunctions of the relative and center-of-mass motions are symmetric with respect to the coordinates of the two nuclei, and the eigenfunction of the rotational motion of the molecular axis is symmetric when the rotational quantum number \( p \) is even, and antisymmetric when \( p \) is odd. With respect to the interchange of the nuclei, therefore, the total wave function of the diatomic molecule of identical nuclei becomes symmetric, if either \( p \) is even and the electron wave function is symmetric, or the former is odd and the latter is antisymmetric. It becomes antisymmetric, on the other hand, if either the former is even and the latter is antisymmetric, or the former is odd and the latter is symmetric.

"For further discussions" Heisenberg "distinguishes" the following two cases. One is the case that the nuclei do not possess any "eigen rotation", and the other is the case that each of the nuclei possesses eigen rotation. If the nucleus is of "specially stable closed structure" like He or O, it seems that the nucleus has no eigen rotation. In this case, because the symmetric and antisymmetric systems of the molecular wave function are not combinative with each other, it turns out that only transitions of the type \( p \rightarrow p \) occur among electron terms of the same kind, but no transitions of the type \( p \rightarrow p \pm 1 \).

Heisenberg supposes that, in the case that "each of the nuclei possesses an eigen rotation, as the electron, of magnitude 1/2", "antisymmetric system can be considered" for the nuclei as for two electrons. As a result, "the nuclear

---

99) M. Born, Zeit. f. Phys. 40 (1926), 167 [the same as Ref. 19) in the present chapter; cf. also the last paragraph of Sec. 2, §1 of the present chapter.]
100) These two limits are nowadays called the separated molecule and the united molecule respectively, making basic concepts in the quantum mechanics of molecules.
magnets” are seen to be antiparallel in the symmetrical systems of the molecular wave function, and parallel in the antisymmetrical systems. The statistical weights of the antiparallel and parallel cases are in the ratio of 3 to 1. Because this ratio is also that between the statistical weights of the symmetrical and antisymmetrical molecular wave functions, the intensity of the rotational spectrum is concluded to vary with the even-oddness of \( p \). Heisenberg’s treatment of the problem under consideration went in this way.

In his paper “On Interpretation of Molecular Spectra II”,\(^{101}\) received on Feb. 7, Hund says that “in part Heisenberg has already shown an explanation” of the various kinds of intensity variation in band spectra. Taking account of the electron spin as well as the angular momentum about the molecular axis, he then studies the spectral terms of a diatomic molecule and analyses, on the basis of the results of his study, the temperature dependence of the rotational specific heat of hydrogen gas. Hund arrives at a conclusion about \( \text{H}_2 \) that the ratio \( \beta \) of the statistical weight of the antisymmetric system of the molecular wave function to that of the symmetric one is nearly \( 1/2 \), and the moment of inertia \( I \) is \( 1.54 \times 10^{-41} \text{ g cm}^2 \).

Hori shows from a detailed experimental study, in his paper “On the Analysis of the Band Spectrum of Hydrogen in Extreme Ultraviolet”\(^{102}\) received on Jul. 19, 1927, that \( I \) is \( 4.67 \times 10^{-41} \text{ g cm}^2 \) and the ratio of the intensity of the antisymmetric term to that of the symmetric term is \( \beta \approx 1/3 \). Citing the manuscript of Hori’s paper, Dennison reanalyses the temperature dependence of the rotational specific heat of hydrogen gas, in his paper “A Note on the Specific Heat of the Hydrogen Molecule”\(^{103}\) received on Jun. 3, 1927, and gets \( I = 4.64 \times 10^{-41} \text{ g cm}^2 \), which is “in substantial agreement” with Hori’s result. “Although it is not possible to fix the value of \( \beta \)”, Dennison finds that “\( \beta = 1/3 \) fits the observed curve distinctly better than \( \beta = 1/3.5 \) or \( \beta = 1/2.7 \)”.

Dennison points out, in a note added in proof on Jun. 16, 1927, that the result \( \beta = 1/3 \) shows that “the nuclear spin is equal to that of the electron”. The spin of the proton is thus known from the study of the symmetric properties of wave functions of diatomic molecules of identical nuclei.\(^{104}\)

---


\(^{102}\) T. Hori, Zeit. f. Phys. \textbf{44} (1927), 834.


\(^{104}\) Until the discovery of the neutron (1932), the nucleus of atomic number \( Z \) and mass number \( A \) was considered to consist of \( A \) protons and \( (A - Z) \) electrons. The finding of the nuclear spin meant thus the constitution of the nucleus by \( (2A - Z) \) particles obeying the Fermi-Dirac statistics, so that the symmetric property for the interchange of two identical
§5.3 Symmetry property of wave function and the binding of hydrogen molecule

Heitler and London showed, in their paper "Interaction of Neutral Atoms and Homopolar Binding according to Quantum Mechanics"\(^{105}\) received on Jun. 30, 1927, that the symmetric property of wave function for the interchange of electrons play an essential role in the binding of a diatomic molecule of identical nuclei. They showed this by a concrete calculation for the H\(_2\) molecule. As we have described in Sec. 2, §1, Chap. 4, Vol. II of the present volume, Lewis and Langmuir analyzed the mode of the functions in chemical binding of electrons and proposed the concept of covalent bonds. Heitler-London's theory gives the essentialistic view of covalent bonds, which provides the basis for the present-day theory of molecules, and is explained in many textbooks on quantum mechanics.

Now, Heitler-London calculate the binding energy of the H\(_2\)-molecule by applying the perturbation method. They assume that, with the interaction between the two atoms neglected, electron 1 belongs to atom \(a\) and is represented by wave function \(\psi_1\), while electron 2 belongs to atom \(b\) and is represented by wave function \(\varphi_2\).\(^{106}\) Though \(\psi_1\varphi_2\) is an unperturbed wave function, it is degenerate with \(\varphi_2\varphi_1\) resulting from the interchange of the two electrons, and "shows a characteristic quantum-mechanical phenomenon of beat, which is closely related to the resonant beat found by Heisenberg".\(^{107}\)

Because \(\psi_1\varphi_2\) and \(\varphi_2\varphi_1\) are independent functions and are not orthogonal to each other, Heitler-London make the linear combinations of \(\psi_1\varphi_2\) and \(\varphi_2\varphi_1\)

\[
\begin{align*}
\alpha &= \frac{\psi_1\varphi_1 + \varphi_2\varphi_1}{\sqrt{2 + 2S}} , \\
\beta &= \frac{\psi_1\varphi_1 - \varphi_2\varphi_1}{\sqrt{2 - 2S}} ,
\end{align*}
\]

(164)

which are orthogonal to each other and normalized, and take them as the unperturbed wave functions. \(S = \int \psi_1\varphi_1\varphi_2\varphi_2 d\tau_1 d\tau_2\) is the integral appearing in the calculation of the normalization constants of these functions. The nuclei should be determined by the even-oddness of \(Z\). This brought various contradictions, which came to be solved only when it was found that the nucleus consists of \(Z\) protons and \((A - Z)\) neutrons, so that the symmetric property under consideration is determined by the even-oddness of \(A\).


\(^{106}\) \(\psi_1\) and \(\varphi_2\): are of the same functional form as a function of the distance between the electron and nucleus of the respective atoms (the normalized wave function of the ground state of a hydrogen atom).

\(^{107}\) Cf. Sec. 1, §2 of the present chapter.
ergy $E_\alpha$ of the $\text{H}_2$ molecule, which is obtained from $\alpha$ by the perturbation calculation, has a minimum value when considered as a function of the distance between the nuclei, while the corresponding energy $E_\beta$ obtained from $\beta$ decreases monotonously with the distance between the nuclei.\(^{108}\) Namely, $\alpha$ gives bound states of the $\text{H}_2$ molecule while $\beta$ gives only a repulsive force for the two H atoms.

Heitler-London say that “it is a result describable only very artificially with classical concepts that two neutral atoms interact with each other in two different ways”. $\alpha$ is symmetric for the interchange of electrons, while $\beta$ is antisymmetric, so that the difference between the symmetry properties of the wave functions discriminates whether the binding takes place or not. This is a completely quantum-mechanical effect. Heitler-London do not treat the spin of the electron. Because any wave function of electrons inclusive of their spin wave functions should be antisymmetric for every interchange of two electrons, $\alpha$ is to be combined with the antisymmetric spin wave function, and $\beta$ with the symmetric spin wave functions. Since the symmetric spin wave function is the eigenfunction of the singlet spin state, it is thus made clear that the covalent bond is effected by two electrons in the singlet spin state.

§6. Formulations of the quantization of wave fields

We have seen before, in §3, i) of the present section, that Dirac proposed the so-called second quantization of Schrödinger’s wave equation of a particle for treating systems of many particles obeying the Bose-Einstein statistics. Dirac noted that the coefficients $b_r$'s in the expansion of the wave function, in terms of the eigenfunctions of the unperturbed Hamiltonian, together with $i\hbar b_r^*$'s form pairs of canonical variables in classical mechanics for the energy of the de Broglie wave. With the normalization $\sum_r |b_r|^2 = 1$, $|b_r|^2$ gives “the probability of the particle being in the state $r$”. With the normalization $\sum_r |b_r|^2 = N$, on the other hand, $|b_r|^2$ means “the probable number of particles in the state $r$”.

$N_r$ and $\theta_r$, which are defined by the relations $b_r = N_r^{1/2} \exp(-i\theta_r/\hbar)$ and $b_r^* = N_r^{1/2} \exp(i\theta_r/\hbar)$, also form a pair of (classical) canonical variables of the problem. Dirac imposed the commutation relation on $b_r$ and $i\hbar b_r^*$ in order to quantize the de Broglie wave. He accordingly put $b_r = (N_r + 1)^{1/2} \exp(-i\theta_r/\hbar)$

\(^{108}\) Heitler-London obtained $E_\alpha$ and $E_\beta$ by estimating some of the integrals appearing in the calculation with the use of their upper limits. The detailed calculation of the integrals was done by Y. Sugiura [Zeit. f. Phys. 45 (1927), 484].
The Establishment of Quantum Mechanics

and \( b^*_r = N^{1/2} \exp(i \theta_r / \hbar) \) to have quantum-mechanical canonical variables \( N_r \) and \( \theta_r \) which satisfy the commutation relation [cf. Eq. (154) in the present chapter]. \( N_r \) has eigenvalues 0, 1, 2, \ldots so that the system of \( N \) particles obeys the Bose-Einstein statistics.

Since for a system of many particles obeying the Fermi-Dirac statistics \( N_r \) can only take eigenvalues 0 and 1, a certain modification of Dirac's method is needed so that it may be applied to this case. The necessary modification was given by Jordan in his paper "On the Quantum Mechanics of the Degeneration of Gas"\(^{109} \) received on Jul. 7, 1927. In this paper, Jordan, referring to his paper\(^{110} \) which was mentioned in §4, ii) of the present section, notices that the property of \( N_r \) having only two eigenvalues of 0 and 1 is similar to that of the spin of the electron. Namely, Jordan considers that \( 2N_r - 1 \) which has two eigenvalues of \( \pm 1 \) can be represented by a matrix similar to that of \( s_z \). Accordingly, \( \theta_r \) is also represented by a matrix.

He puts

\[
N_r = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \theta_r = \frac{\hbar}{4} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},
\]

and defines \( b_r \) and \( b^*_r \) as

\[
\begin{align*}
b_r &= \exp(-\theta_r / \hbar) N_r^{1/2} = (1 - N_r)^{1/2} \exp(-i \theta_r / \hbar) \\
b^*_r &= N_r^{1/2} \exp(i \theta_r / \hbar) = (i \theta_r / \hbar) N_r^{1/2},
\end{align*}
\]

to show that the relations needed for the Fermi-Dirac statistics, \( b^*_r b_r = N_r \), \( N_r^2 = N_r \) and \( (1 - N_r)^2 = 1 - N_r \), are satisfied. The Hamiltonian of the system of \( N \) particles is given by \( F = \sum_{rs} b^*_r H_{rs} b_s \) as in Dirac's paper \([H_{rs} \text{ being equal to } W_{rs} \delta_{rs} + V_{rs} \text{ in Eq. (150) given in §3.1 of the present section}].\]

Because of the relation \( \exp(i \theta_r / \hbar) = (2i / \pi \hbar) \theta_r \), Jordan writes the Schrödinger equation for the system of \( N \) particles "symbolically" as

\[
\left\{ \sum_{rs} H_{rs} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_r \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_s + \frac{\hbar}{2\pi i} \frac{\partial}{\partial t} \right\} \Psi = 0.
\]

\(^{109} \)P. Jordan, Zeit. f. Phys. 44 (1927), 473.

\(^{110} \)P. Jordan, Zeit. f. Phys. 44 (1927), 1 [the same as Ref. 91] in the present chapter.
Jordan shows also that the fluctuation in density of a system, calculated in the case of a string as in Born-Heisenberg-Jordan's paper, is proportional to \[ \sum_r N_r (1 + N_r) \] using the Bose-Einstein statistics, and to \[ \sum_r N_r (1 - N_r) \] using the Fermi-Dirac statistics. Furthermore, Jordan in a paper worked out in collaboration with Wigner, "On Pauli's Prohibition of Equivalents", received on Jan. 26, 1928, treats a gas of molecules obeying the Fermi-Dirac statistics. They show that the quantization by imposing "the particular noncommutable properties of multiplication" on "the amplitude", given in terms of \( b_r \)'s of "a three-dimensional wave field", that is, of the de Broglie wave field, leads to "the existence of corpuscular gas atoms", and at the same time to "the validity of Pauli's prohibition of equivalents".

"The particular noncommutable properties of multiplication" are identical with what are nowadays called anticommutation relations. They are given by

\[
b_r^* b_s + b_s b_r^* = \delta_{rs},
\]

and result from the fact that \( b_r \) and \( b_r^* \) in Eq. (166) satisfy the relations \( b_r^* b_r = N_r \) and \( b_r b_r^* = 1 - N_r \). Thus, it is made clear that the system of many particles obeying the Fermi-Dirac statistics can be treated by means of the second quantization of the de Broglie wave field with the use of anticommutation relations.

Dirac's and Jordan's methods of the second quantization mentioned above are both related to ideal gases. On the other hand, Jordan presents a system of charged particles obeying the Bose-Einstein statistics "by a 'quantization' of the de Broglie wave", in a paper collaborated with Klein "On the Many Body Problem in Quantum Mechanics" received on Oct. 4, 1927. Jordan-Klein put the wave equation of the de Broglie wave in the form

\[
\left( -\frac{\hbar^2}{2m} \Delta \phi + \frac{\hbar}{i} \frac{\partial \phi}{\partial t} - \epsilon V \phi \right) = 0,
\]

where \(-\epsilon\) is the electric charge of the electron, and \(V(r, t)\) is the electric potential which is determined by the Poisson equation

\[
\Delta V = 4\pi \epsilon \phi^* \phi,
\]

---

111) Cf. Sec. 2, §6, Chap. 1 of the present volume.
114) Jordan-Klein do not use the notation \( \hbar (= \hbar/2\pi) \), but we use here it for convenience.
given in terms of the density of electric charge \(-\varepsilon \varphi^* \varphi\). With the use of the Green function \(G(r, r') = 1/|r - r'|\), the electric potential is given as a function of \(r\) in the form

\[
V(r) = V_0(r) - \varepsilon \int G(r, r') \varphi^*(r') \varphi(r') d\tau',
\]

(171)

where \(V_0(r)\) is the electric potential of an external electric field.

These equations can be derived from the Hamilton principle \(\delta \int_{t_0}^t dt \int L d\tau = 0\) in classical dynamics if the Lagrangian density \(L\) is taken to be

\[
L = \frac{\hbar^2}{2m} (\nabla \varphi^* \cdot \nabla \varphi) + \frac{\hbar}{i} \left( \varphi^* \frac{\partial \varphi}{\partial t} - \frac{\partial \varphi^*}{\partial t} \varphi \right) - \varepsilon V \varphi^* \varphi - \frac{1}{8\pi} (\nabla (V - V_0))^2.
\]

(172)

The energy of the wave field \(F\) is given by

\[
F = \int \left\{ \frac{\hbar^2}{2m} (\nabla \varphi^* \cdot \nabla \varphi) - \varepsilon V_0 \varphi^* \varphi + \frac{1}{8\pi} (\nabla (V - V_0))^2 \right\} d\tau.
\]

(173)

If \(u_s\) is the normalized solution of the eigenvalue equation

\[-(\hbar^2/2m)\Delta u_s - \varepsilon V_0 u_s = E_s u_s,
\]

(174)

\(E_s\) being the eigenvalue, and if \(\varphi\) is developed as

\[
\varphi(r, t) = \sum_s b_s(t) u_s(r),
\]

(175)

\(F\) is expressed as

\[
\left\{ \begin{aligned}
F &= \sum_s b_s^* b_s E_s + \left( \frac{\varepsilon^2}{2} \right) \sum_{\mu \nu rs} A(\mu \nu / rs) b_r^* b_s^* b_\mu b_\nu , \\
A(\mu \nu / rs) &= \int \int G(r', r'') u_\mu(r') u_\nu(r'') u_r(r') u_s(r'') d\tau' d\tau''
\end{aligned} \right.
\]

(176)

where the \(u_s\)'s are considered to be real functions.

Since the Hamilton principle gives the relations \((\hbar/i) b_s = -\partial F/\partial b_s^*\) and \((\hbar/i) b_s^* = \partial F/\partial b_s\) so that \((\hbar/i) b_s\) and \(b_s^*\) form a pair of canonical variables, Jordan-Klein impose, following Dirac, the commutation relations \(b_r b_s^* - b_s^* b_r = \delta_{rs}\), and introduce variables \(N_r\) and \(\theta_r\) to express Eq. (156) for the wave function \(\Psi(N_1, N_2, \ldots)\), given in §3.1 of the present section, with \(F\) given by Eq. (176).

In Dirac’s quantum theory of the interaction of an atom and electromagnetic field, the method of quantization was not directly applied to the latter.
The quantization of the electromagnetic field is treated in Jordan and Pauli's paper "On the Quantum Electrodynamics of Charge-free Fields", received on Dec. 7, 1927. Jordan-Pauli first make the Fourier development of the free electromagnetic field in a cubic cavity of length $L$ and quantize it by imposing commutation relations on the coefficients in the development taken as the canonical variables, and then take the limit $L \to \infty$ in the results so obtained. The method used by them is the same as what one sees in present-day textbooks on quantum mechanics.

They introduce the propagation function of an electromagnetic field, $\Delta(x, y, z, t)$, in a form invariant with respect to the Lorentz transformation. Using it they derive the commutation relations of electromagnetic fields at two space-time points $P(x, y, z, t)$ and $P(x', y', z', t')$ themselves, $[E_x(P), H_y(P')]$ and so on, and systematize them concisely by means of the 4-electromagnetic potential. When the electromagnetic field is quantized by the commutation relations for the electromagnetic fields themselves, the wave function $\Psi$ changes from a simple function to a functional. Jordan-Pauli then discuss the formulation necessitated by this change.

Though they treat only free electromagnetic fields, their way of consideration "can completely be applied also to material waves of freely moving particles, if one has to do with particles of the same kind which obey the Einstein-Bose statistics". However, "in the other case of particles with the Fermi statistics, the quantization of the material waves is not yet completely clarified". For nonrelativistic material waves, as is mentioned above, Jordan-Wigner present, soon after, quantization by means of the anticommutation relation. Jordan-Pauli say that "concerning a general relativistically invariant quantum theory still to be found for the wave fields, which on the one hand will have to consider also the electromagnetic fields corresponding to the existence of charged particles, and on the other hand will have to take into account the influence of the electromagnetic waves on the material waves, it may well be expected that such a theory will include in itself commutation relations of the free electromagnetic radiation fields here established, as well as those of the material waves of free particles as special limiting cases".

---

116) The propagation function $\Delta(x, y, z, t)$ is the Green function of the wave equation of the electromagnetic field, and describes the aspect of the propagation of electromagnetic waves.
117) $[A, B] = AB - BA$. 
Jordan-Pauli's expectation comes to be confirmed by Heisenberg and Pauli in their paper "On the Quantum Dynamics of Wave Fields",\textsuperscript{118} received on Mar. 19 following Dirac's presentation of the relativistic wave equation of the electron. Taking the Dirac relativistic electron wave field and the electromagnetic field generated by it, Heisenberg-Pauli divide the space containing the fields into small cells of volume $\Delta V$ to express the Lagrangian, and impose the commutation relations on the strengths of the fields in the cells at every time being taken to be the canonical variables. They then consider the limit of $\Delta V \to 0$, arriving at the quantization of the interacting electron and electromagnetic fields. They show also that this methods gives the commutation relations of the electromagnetic fields derived by Jordan-Pauli with the use of the propagation function. In this way, the formulation of the quantum theory of fields is fundamentally established.

§7. The Dirac equation and the Klein-Nishina formula

We have described in §4.2 of the present section that Pauli's theory of spin was "provisional and approximate", as was noted by Pauli himself. In his theory, it was presumed that the $z$-component, $s_z$, of the spin takes only two values and that the coefficient of magnetic dipole moment of the spin is twice the usual value. This theory was not put in a form of relativistic invariance, too. Furthermore, as was pointed out by Jordan, $s_z$ and its canonical conjugate are of such a character that the general principle of quantum mechanics could not be applied to them.

Dirac tackles these problems in his paper "The Quantum Theory of the Electron"\textsuperscript{119} received on Jan. 2, 1928, presenting the relativistic wave equation of the electron. Dirac considers the "duplexity phenomena" and asks the question "why Nature should have chosen this particular model for the electron instead of being satisfied with the point-charge". He analyses the difficulties based on the thought that "one would like to find some incompleteness in the previous methods of applying quantum mechanics to the point-charge electron such that, when removed, the whole of the duplexity phenomena follow without arbitrary assumptions", and shows that "the incompleteness of the previous theories lies in their disagreement with relativity, or, alternatively, with the general transformation theory of quantum mechanics".

\textsuperscript{118}W. Heisenberg und W. Pauli, Zeit. f. Phys. \textbf{56} (1929), 1.
Citing Gordon’s and Klein’s papers which we have mentioned in §2.1 of the present section, Dirac points out the difficulty that, though their relativistic equation can give the emission and absorption of radiation, if $j_\alpha$ given in Eq. (145) is interpreted to be the four-current density, it cannot give the probability density of any dynamical variable other than the position of the electron. Because the space components $j = (j_1, j_2, j_3)$ of $j_\alpha$ given by Eq. (145) and $\rho$ determined from its time-components $j_4$ through the relation $j_4 = ic\rho$ satisfy the continuity equation $\partial \rho / \partial t + \text{div} \, j = 0$, Gordon and Klein supposed $\rho$ and $j$ to be the charge and current densities respectively.

In the case of the nonrelativistic Schrödinger wave equation, a continuity equation of this kind is also satisfied by $\rho = \bar{\psi}\psi$. For this reason Schrödinger assumed $e\bar{\psi}\psi$ to be the charge density and $\bar{\psi}\psi$ “the weight function”, as has been described in Sec. 5, §1 of the last chapter. With the establishment of quantum mechanics, $\bar{\psi}\psi$ was interpreted as the probability density and $\bar{\psi}Q\psi$ as the probability distribution of a dynamical quantity $Q$. Dirac means that, in the case of the relativistic wave equation of the Klein-Gordon type, it is not clear how the probability distribution of the momentum or any other dynamical variable should be defined, even if the probability density is defined “by the use of $\rho$”, because $\rho$ has a term proportional to $\{\bar{\psi}(\partial \psi / \partial t) - \psi(\partial \bar{\psi} / \partial t)\}$.

Dirac says that the interpretation in the nonrelativistic quantum mechanics mentioned above “is made possible” by the wave equation being linear in the energy $W$ or $i\hbar \partial / \partial t$, so that “the wave function at any time determines the wave function at any later time”.\textsuperscript{120}\textsuperscript{121} “The wave equation of the relativity theory must also be linear in $W$”. In the absence of electromagnetic field, the wave equation of Klein-Gordon type is

$$(-p_0^2 + p^2 + m^2c^2)\psi = 0; \quad p_0 = W/c = (i\hbar/c)\partial / \partial t.$$ \hfill (177)

“The symmetry between $p_0$ and $(p_1, p_2, p_3)$ required by relativity shows that, since the Hamiltonian we want is linear in $p_0$, it must also be linear in $(p_1, p_2, p_3)$”.

Thus, Dirac assumes that “our wave equation” is of the form

$$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)\psi = 0, \quad (178)$$

\textsuperscript{120}Cf. Sec. 2, §1, p. 160 of the present chapter.

\textsuperscript{121}As Pauli pointed out in Part I “General Principles of Wave Mechanics”, Vol. 24 of “Hanbuch der Physik” (1933), in the case of an equation of Klein-Gordon type one cannot assume that $\{\bar{\psi}(\partial \psi / \partial t) - \psi(\partial \bar{\psi} / \partial t)\} \geq 0$ at any arbitrary time, because the values of $\psi$ and $\partial \psi / \partial t$ can be chosen independently.
where "the dynamical variables or operators $\alpha_1, \alpha_2, \alpha_3, \beta$ are independent of $p_0, p_1, p_2, p_3$, i.e., commute with $t, x_1, x_2, x_3"$. Equation (178) becomes, when operated on by $(-p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3)$, an equation consisting of $p_0^2$ and a homogenous quadratic form of $p_1, p_2, p_3, \beta$. Dirac considers that the resultant equation should agree with Eq. (177), and sets up the conditions for this agreement, which are expressed as $\alpha_\mu^2 = 1$, $\alpha_\mu \alpha_\nu + \alpha_\nu \alpha_\mu = 0 (\mu \neq \nu)$, $\mu, \nu = 1, 2, 3, 4$, if $\beta$ is put to be $\alpha_4 mc$.

If $\mu$ and $\nu$ are restricted to 1, 2, 3, these relations can be satisfied by the Pauli spin matrices $s_1, s_2, s_3$ given by Eq. (162). However, $s_1, s_2, s_3$ cannot be taken to be $\alpha_1, \alpha_2, \alpha_3$. This is "because then it would not be possible to find the fourth", $\alpha_4$. Thus, Dirac "extends s's in a diagonal manner", and introduces the four by four matrices

$$\begin{pmatrix} s_r & 0 \\ 0 & s_r \end{pmatrix} \quad (r = 1, 2, 3)$$

$$\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad \rho_4 = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix},$$

showing that

$$\alpha_r = \rho_1 \sigma_r \quad (r = 1, 2, 3), \quad \alpha_4 = \rho_3 \quad (179)$$

satisfy the conditions mentioned above. Consequently, the wave function $\psi$ is to have four components. With $p_4 = -ip_0$ and $\gamma_r = \rho_2 \sigma_r$, $\gamma_4 = p_3$, that is,

$$\gamma_r = i \begin{pmatrix} 0 & -s_r \\ s_r & 0 \end{pmatrix} \quad (r = 1, 2, 3), \quad \gamma_4 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad (181)$$

the wave equation (178) can be written in the form

$$(i \sum_\mu \gamma_\mu p_\mu + mc) \psi = 0, \quad (\mu = 1, 2, 3, 4), \quad (182)$$

which is more symmetrical in $p_1, p_2, p_3, p_4$. $\gamma$'s satisfy the relation $\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta_{\mu\nu}$.

In order to show the relativistic invariance of the wave equation so obtained, Dirac transforms $p_\mu$ into $p'_\mu = \sum_\nu a_{\mu\nu} p_\nu$ by means of the Lorentz

---

122) $0$ and $I$ in Eqs. (179) and (181) are the two-by-two zero and unit matrices respectively.
transformation. Equation (182) is then transformed into

\[
(i \sum \gamma'_\mu p'_\mu + mc) \psi = 0, 
\]

where \(\gamma'_\mu\) is defined as \(\gamma'_\mu = \sum_\nu a_{\mu\nu} \gamma_\nu\). From the properties of \(\alpha_{\mu\nu}\) it is derived that \(\gamma'_\mu \gamma'_\nu + \gamma'_\nu \gamma'_\mu = 2 \delta_{\mu\nu}\), which is identical with the relation satisfied by the \(\gamma_\mu\)'s. With the use of this relation and the general properties of four-by-four matrices, Dirac shows that the \(\gamma'_\mu\)'s can be put into the same form as Eq. (181) by means of a canonical transformation, proving the relativistic invariance of the Dirac wave equation (182) or (178).

In the case of an electron in an electromagnetic field, Dirac substitutes \(p'_\mu + (e/c)A_\mu\) for \(p_\mu\), \(A_\mu\) denoting the 4-potential. He then deduces that, as a result of the Dirac equation, the sum of the angular momentum and \(h(\sigma/2)\) of an electron in a spherically symmetric static electric field is a constant of the motion,\(^{123}\) the spin angular momentum \(h(\sigma/2)\) possesses the anomalous coefficient 2 of magnetic dipole moment, and the so-called spin-orbit interaction in a central force field is of the required strength.\(^{124}\) In this way, Pauli’s “provisional and approximate” theory of the spin is elevated, by means of the extension “in a diagonal manner” of Pauli’s spin matrices, to Dirac’s essential theory.

In Part II\(^{125}\) of his paper received on Feb. 2, 1928, Dirac shows that, if one defines \(\rho\) and \(j_r\) in terms of \(\psi\) satisfying the wave equation for an arbitrary electromagnetic field and \(\phi\) satisfying the wave equation obtained by transposing all the matrices in the wave equation for \(\psi\) as

\[
\rho = \phi \gamma_0 \chi, \quad j_r = -\phi \gamma_r \chi \; (r = 1, 2, 3),
\]

where \(\gamma_0 = \rho_3\) and \(\chi = \gamma_0 \psi\), then \((\rho, j)\) form a 4-vector and satisfy the equation of continuity \(\partial \rho / \partial (ct) + \text{div} \, j = 0\). From this fact, Dirac suggests that \(\rho\) gives the probability density of the electron, following Schrödinger-Born’s thought.

Klein and Nishina apply the Dirac equation to the Compton scattering in their paper “On the Scattering of Light by Free Electrons according to

\(^{123}\) \(\sigma\) is a vector with components \((\sigma_1, \sigma_2, \sigma_3)\).

\(^{124}\) Dirac derived the spin-orbit interaction by means of a perturbation method. It was done by C. G. Darwin [Proc. Roy. Soc. A118 (1928), 654] and W. Gordon [Zeit. f. Phys. 48 (1928), 11] to solve exactly the Dirac wave equation for the hydrogen atom, and derive the fine structure of its energy levels consistent with experimental results.

Dirac’s New Relativistic Quantum Dynamics”\textsuperscript{126} received on Oct. 30, 1928. Their method of calculation is a correspondence-theoretical one similar to that in Gordon’s paper which we have mentioned in §2.2 of the present section. Namely, with the use of the vector potential of the incident plane wave, Klein-Nishina solve, in the form of a Fourier expansion, the Dirac equation, instead of the wave equation of Klein-Gordon as in Gordon’s case, and calculate the vector potential of the radiation generated by the source $-e_j$ determined by the solution, to obtain the intensity of the scattered radiation.\textsuperscript{127}

Before the paper under consideration, Klein-Nishina wrote a letter\textsuperscript{128} to the Editor of Nature, in which they compared their result with Compton’s experimental result\textsuperscript{129} for $\gamma$-rays from RaC (wavelength 0.022Å, $\alpha = h\nu/mc = 1.1$), as in Dirac’s paper mentioned in §2.2 of the present section. Their value of the intensity of the scattered radiation is larger than Dirac’s with a maximum deviation of 82.5\% (at 125°), while the experimental values are less than Dirac’s “in roughly the same ratio (75%)” at all the angles as noted by Dirac. Also, their value of the ratio of the absorption coefficient $\sigma$ to that in the classical theory, $\sigma_0$, becomes $\sigma/\sigma_0 = 1 - 2\alpha + 26\alpha^2/5$ for $\alpha \ll 1$ and $\sigma/\sigma_0 = (3/8\alpha)(\log 2\alpha + 1/2)$ for $\alpha \gg 1$, and is larger by about 50\% than Dirac’s at $\alpha = 1$. Klein-Nishina add a note in the proof, in which they say that from recent experiments the $\gamma$-rays from RaC would appear very complex with an average wavelength of only about half the value of 0.022Å, and that if this is taken into account their values of the intensity of the scattered radiation is “found to agree with the measurements rather better”. They point out also that if their formula for absorption is used, “the wavelengths obtained for the cosmic rays are considerably shorter than those ordinarily assumed”.

In Klein-Nishina’s paper, the formulation of the quantization of wave fields is not applied to electromagnetic fields and the Dirac electron field. This formulation is presented in Heisenberg-Pauli’s paper received about five months later than Klein-Nishina’s, as we have mentioned at the end of §6 of the present section. Further about one year later, Tamm calculates the Compton effect in his paper “On the Interaction of Free Electron with Radiation according to

\begin{itemize}
\item[]\textsuperscript{126}O. Klein und Y. Nishina, Zeit. f. Phys. 52 (1928), 853.
\item[]\textsuperscript{127}Nishina calculated, in his paper received on the same day as Klein-Nishina’s [Zeit. f. Phys. 52 (1929), 869], the polarization of the scattered radiation.
\item[]\textsuperscript{128}O. Klein and Y. Nishina, Nature 122 (Sep. 15, 1928), 398.
\item[]\textsuperscript{129}A. H. Compton, Phys. Rev. 21 (1923), 483 [the same as Ref. 6) in Sec. 1, §2, Chap. 5, Vol. II of the present book].
\end{itemize}
the Dirac Theory of Electron and according to Quantum Electrodynamics"\textsuperscript{130} received on Apr. 7, 1930, by means of Heisenberg-Pauli's formulation of the quantum field theory, saying that Klein-Nishina's "correspondence-theoretical treatment" is "not only logically unsatisfactory but also ambiguous", and arrives at the same result as Klein-Nishina's. The Klein-Nishina scattering formula has been confirmed by later experiments, thus becoming an important formula for the treatment of elementary processes in the interaction of radiation with matter.

Now, Dirac points out in the first paper the second difficulty in the wave equation of Klein-Gordon type in the case of an arbitrary electromagnetic field, besides the first one mentioned above. It is the difficulty that the complex conjugate of this wave equation is identical with that obtained by replacing the electric charge of the electron $-e$ with $e$. "The wave equation thus refers equally well to an electron with charge $e$ as to one with charge $-e$". In the latter case, "the energy $W$ has a negative value". One gets over the difficulty on the classical theory "by arbitrarily excluding" such solutions, but "one cannot do this on the quantum theory, since in general a perturbation will cause transitions from states with $W$ positive to states with $W$ negative".

Dirac says that in his paper under consideration he is "concerned only with the removal of the first of these two difficulties".

About the second difficulty still remaining in the Dirac equation, Dirac supposes in 1930 that "all the states of negative energy are occupied except perhaps a few of small velocity", so that transitions of the kind mentioned above are forbidden by the Pauli exclusion principle. He considers then that if there arise vacant states or "holes" in the distribution of negative-energy electrons, these holes "will be things of positive energy" and will therefore be like ordinary particles of positive charge, and assumes that the holes are the "protons". After some analyses made by H. Weyl, Ig, Tamm, J. R. Oppenheimer and others, Dirac presents in 1931 the theory of positrons in which the holes are not protons but the "anti-electrons". Though this process is interesting enough, it is not a matter of the history of quantum mechanics, but of the history of the theory of elementary particles.

It can be said that the Dirac wave equation is obtained by making the Schrödinger nonrelativistic equation of the de Broglie wave (a partial differential equation of first order in time) relativistic. The so-called quantization

\textsuperscript{130}Ig. Tamm, Zeit. f. Phys. 82 (1930), 545.
of the Dirac wave equation containing $\hbar$, that is, the application of the quantum field theory to the Dirac relativistic electron field, is done in 1929 by Heisenberg-Pauli, as we have described in §6 of the present section. By this procedure, $\rho$ given in Eq. (183) becomes the density of the electron field, instead of the probability density in Schrödinger-Born’s sense, and the wave function $\psi$ an operator to act on the state vector $\Psi$ of the electron field, instead of the probability amplitude, $|\Psi|^2$ giving the probability of the state of the electron field which is indicated by the number of electrons and by the momentum, energy and spin of each electron.

In the Heisenberg-Pauli theory, the electron field is considered as the source of electromagnetic field, too. After the proposal of the Dirac positron theory, the process in which an electron in a negative-energy state absorbs a photon to transit into a positive-energy state comes to be grasped as the creation of a pair of electron and positron by a photon, and its inverse process as the annihilation of a pair of electron and positron accompanied by the creation of a photon. However, the characteristic is still maintained that the electron field is the source of electromagnetic field. Similar situation comes about for the nucleon field and the meson field, when the Yukawa meson theory is presented in 1935 and the quantum field theory is applied to it. These situations are seen to be the manifestation of “the conflict between matter and field”,¹³¹ and the solutions of this conflict still remains one of the subjects of the present-day theory of elementary particles.¹³²

In this sense, quantum mechanics is completed with the Dirac relativistic wave equation. Heisenberg-Pauli’s quantum theory of relativistic fields and Dirac’s theory of positron mediate quantum mechanics so completed to the theory of elementary particles, a new stage in the cognition of Nature.

Chapter 4

The Logic of Quantum Mechanics

1. Uncertainty "Principle" and Complementarity "Principle"

§1. Heisenberg's uncertainty relation

As we have seen in the last chapter, matrix mechanics and wave mechanics were unified, by means of Born's idea of grasping the wave function as the probability amplitude and Dirac-Jordan's transformation theory of understanding the wave function as the matrix element of the transformation from the eigen vector of a physical quantity to that of another, so that the state of an atomic system came to be represented by the state vector in the Hilbert space. As we have described also in the previous chapters, matrix mechanics was a phenomenological theory while wave mechanics was a substantialistic theory. Because quantum mechanics was established through the unification of these two, there should be included in quantum mechanics certain logic characteristic of this unification.

Though the logic proper to quantum mechanics was not immediately grasped clearly and expressed explicitly by physicists, attempts were made to understand the various aspects which are taken on by quantum mechanics in ways qualitatively different from classical mechanics. Heisenberg's idea of the so-called uncertainty "principle" and Bohr's view of the so-called complementarity "principle" were presented as such attempts.

Heisenberg stresses, in his paper "On the Intuitive Content of the Quantum Theoretical Kinematics and Dynamics"\(^1\) received on Mar. 23, 1927, that in quantum mechanics "canonically conjugate quantities can be simultaneously

\(^1\)W. Heisenberg, Zeit. f. Phys. 43 (1927), 172.
determined only with a characteristic inaccuracy", and "this inaccuracy is the original ground for the occurrence of statistical relations in quantum mechanics".

"It is our belief that we understand a physical theory to be intuitive, if we can ourselves think the experimental consequences of this theory qualitatively in all simple cases, and if we have at the same time recognized that the application of the theory never contains inner contradictions". Heisenberg's paper cited here which begins with this sentence is a long one of twenty-nine pages written under Bohr at Copenhagen. In this paper, Heisenberg derives from a "thought experiment" the "inaccuracy" mentioned above, that is, the uncertainty relation, explains it by means of Dirac-Jordan's transformation theory, and treats "qualitatively" certain quantum-mechanical phenomena to interpret the quantum-mechanical laws, arriving at the supposition of the negation of causality in quantum mechanics. We analyze here in particular the line of Heisenberg's thinking on the probability concept and causality in quantum mechanics.

Heisenberg says that "the intuitive interpretation of quantum mechanics is until now still full of inner contradictions, which result in the struggle of the views about discontinuous and continuous theories, particles and waves". He stresses that "even from this one would infer that any interpretation of quantum mechanics with the usual kinematic and mechanical concepts is any way not possible", but "quantum mechanics was originated just from the attempt to break the usual kinematic concepts and to set, instead of them, relations between concrete experimentally given numbers".

This means that matrix mechanics began with Heisenberg's work to suppose the existence of "aggregates" of quantum-theoretical amplitudes corresponding to the amplitudes in classical mechanics of electrons in an atom, and to search for the relations between them by taking Ritz's combination rule and Kramers-Heisenberg's dispersion theory as clues [cf. Sec. 1, Chap. 1 of the present volume]. Schrödinger criticized Heisenberg's phenomenological view as Machismic, and showed, from the substantialistic viewpoint of treating the de Broglie wave "in so intuitionally concrete form as of something quite real", that the matrix elements were determined by wave functions.

However, because of the nature of the wave function $\psi$ as a function in a multi-dimensional configuration space, Schrödinger was obliged to get out of the substantialistic image, arriving at the reinterpretation that $|\psi|^2$ should be the "weight function". On the other hand, Born supposed that $|\psi|^2$ should
be the probability density of finding the system in a given state, using as a
guide the idea that the de Broglie wave would guide the energy transfer, and
thus make possible the unification of wave and quantum, an idea similar to
de Broglie’s [cf. Secs. 3 and 5, Chap. 2 of the present volume].

Now, Heisenberg thinks that because “it seems that this has succeeded”,
namely, relations between experimentally obtained numbers have been set
up, “the mathematical scheme of quantum mechanics will need no revision”.
In quantum mechanics as such, however, “there should exist the relation
$\mathbf{pq - qp = h/2\pi i}$”.\(^2\) He thinks that “we have therefore good reason to become
suspicious of the uncritical application of the words ‘position’ and ‘velocity’”,
and puts the question “whether it is possible, by means of a more precise anal­
ysis of the kinematical and mechanical concepts, to resolve the contradictions
existing until now in the intuitive interpretation of quantum mechanics and to
arrive at an intuitive understanding of the quantum mechanical relations”.

Heisenberg says that if one wants to be clear about “what should be under­
stood” by the word “the position of the object”, for example of the electron,
“one must state definite experiments, with the help of which one intends to
measure the ‘position of the electron’”. As an example of such experiment, he
considers the observation of the electron under a microscope. For the reason
that “the highest reachable accuracy of the determination of position is given
here essentially by the wavelength of the light used”, Heisenberg proposes the
famous thought experiment of $\gamma$-ray microscope, by considering that “one will
be able in principle to construct for instance a $\gamma$-ray microscope and to carry
out with it the determination of position as accurately as one wants”. In this
determination, however, “there is essentially an accompanying circumstance:
the Compton effect”. A light quantum hits the electron and is reflected, and
then on being deflected by the lens of the microscope causes the photoelectric
effect which is observed by us. Heisenberg says that, because “at the moment
when the light quantum is reflected by the electron, the electron changes its
momentum discontinuously”, “at the moment when the position of the electron
is known, its moment can be known only down to quantities which correspond
to those discontinuous changes”.

Namely, letting $q_1$ be “the accuracy (for example the average error)” with
which the position coordinate $q$ is known, “therefore here the wavelength of the
light”, and $p_1$ be “the accuracy” with which the momentum $p$ is determined,

\(^2\) Gothic letters are used here to denote matrices.
Heisenberg considers that the relation
\[ p_1 q_1 \sim h \] (1)
holds on account of the formula for the Compton effect. This relation is later called the Heisenberg uncertainty relation.

He says that a similar relation holds for the energy \( E \) and the time \( t \), or for the action variable \( J \) and the angle variable \( w \), by noting the following. In the Stern-Gerlach experiment the magnetic dipole moment of an atom is determined, while the magnetic dipole moment depends only on \( J \) of the atom, which is related to the energy of a stationary state of the atom. The Stern-Gerlach experiment can therefore be applied to the measurement of the energy of the stationary state of the atom. If \( E_1 \) is “the accuracy” of this measurement and \( d \) the linear size of the beam of atoms passing through the aperture, the force deflecting the beam is at most of the order of \( E_1/d \). The deflection angle of the beam is then given by \( E_1 t_1/pd \), where \( p \) is the momentum of the atom and \( t_1 \) the time during which the beam is under the action of the deflecting force. The deflection angle is, on the other hand, expected to be of the same order of magnitude as the diffraction angle of the beam at the aperture, which is given by \( \lambda/d \) in terms of the de Broglie wave length \( \lambda \). With the use of \( \lambda = h/p \), one gets thus the relation
\[ E_1 t_1 \sim h, \] (2)
which gives \( J_1 w_1 \sim h \) for \( J \) and \( w \).

§2. The uncertainty relation and commutation relation

Heisenberg points out that the relation (1) is “the precise expression” for the fact that “earlier”, namely in the old quantum mechanics, the phase space was divided into cells of size \( h \), and also shows that the relation (1) is “in direct mathematical connection with the commutation relation”, by using the probability amplitude as formulated generally by Dirac and Jordan in the way we have described in Sec. 2, §4 ~ §6 of the last chapter, that is, the transformation function.

If the value \( q' \) of \( q \) is determined with “the accuracy” \( q_1 \) when the value of a certain parameter is definitely determined to be \( \eta \), the probability amplitude \( S(\eta, q) \) differs marked by from zero only in a region of \( q \) around \( q' \) having a size of the order of magnitude \( q_1 \). Thus, Heisenberg assumes as the form of
$S(\eta, q)$ in this case

$$S(\eta, q) \propto \exp \left\{ -\frac{(q - q')^2}{2q_1^2} - \frac{2\pi i p'(q - q')}{h} \right\}. \quad (3)$$

$|S(\eta, q)|^2$ given by this form is proportional to $\exp\{- (q - q')^2/q_1^2\}$. According to Dirac-Jordan's formulation, the probability amplitude with respect to $p$ is given by

$$S(\eta, p) = \int S(\eta, q)S(q, p)dq. \quad (4)$$

Putting

$$S(q, p) = \exp \left( \frac{2\pi ipq}{h} \right) \quad (5)$$

after Jordan [cf. Eq. (103) of the last chapter], one gets

$$S(\eta, p) \propto \exp \left\{ -\frac{(p - p')^2}{2p_1^2} - \frac{2\pi ip'(p - p')}{h} \right\}, \quad (6)$$

where $p_1$ is defined by

$$p_1q_1 = \frac{h}{2\pi}. \quad (7)$$

Because $|S(\eta, p)|^2$ is proportional to $\exp\{- (p - p')^2/p_1^2\}$, it is seen that $p_1$ and $q_1$ given here satisfy the relation (1).

In Heisenberg's derivation of the uncertainty relation, the commutation relation is not directly applied but is shown to be "in direct mathematical connection with the commutation relation" on a concrete and appropriate assumption for the form of $S(\eta, p)$, by virtue of the equivalence of $p$ to $-i\hbar \partial/\partial q$ with respect to $S(q, p)$ in Dirac-Jordan's formulation.

The direct derivation of the uncertainty relation from the commutation relation is performed by Kennard,\textsuperscript{3} Weyl,\textsuperscript{4} Condon,\textsuperscript{5} Robertson\textsuperscript{6} and others, and is finally given in the form as is seen in present-day text books on quantum

\textsuperscript{3}E. H. Kennard, Zeit. f. Phys. 44 (1927), 326.
\textsuperscript{4}H. Weyl, "Gruppentheorie und Quantenmechanik" (S. Hirzel, Leipzig, 1928).
\textsuperscript{5}E. U. Condon, Science 64 (1929), 573.
\textsuperscript{6}H. P. Robertson, Phys. Rev. 34 (1929), 163.
mechanics. Namely, as is well known, if the mean value $A_0$ of a physical quantity represented by the operator $A$ is defined by

$$A_0 = \int \bar{\psi} A \psi d\tau,$$

(8)

and the mean squared deviation $(\Delta A)^2$ by

$$(\Delta A)^2 = \int \bar{\psi} (A - A_0)^2 \psi d\tau,$$

(9)

the integration being over the whole space of coordinates, it can be shown by the use of the Schwarz inequality that, when $A$ obeys the commutation relation

$$AB - BA = \left(\frac{\hbar}{2\pi i}\right) C$$

(10)

together with operators $B$ and $C$ representing other physical quantities, there holds the relation

$$\langle \Delta A \Delta B \rangle \geq \left(\frac{\hbar}{4\pi}\right) |C_0|.$$  

(11)

In connection with the uncertainty relation (1), Heisenberg derives the spreading of a “wave packet” with the use of the transformation function. Namely, he calculates from $S(\eta, q)$ at time $t = 0$ as given by Eq. (3), $S(\eta, q)$ at the time $t$, by means of the transformation function between the matrix representation in which $q$ is diagonal at $t = 0$ and that in which $q$ is diagonal at $t$, showing that the “inaccuracy” $q_1$ of $q$ spreads to $(q_1^2 + \hbar t/2\pi m)^{1/2}$, where $m$ is the mass of the particle. Also, Heisenberg discusses the scattering of electrons from a crystal lattice from the viewpoint of the uncertainty relation (1), and treats, in relation with the uncertainty relation (2), the problems of the time of the transition of an atom, the resonant fluorescence of an atom irradiated by a light, etc.

7) Heisenberg considers that scattering into the same specified directions as in the Bragg reflection occurs when the wavelength of the electrons is of the order of the lattice constant, but the classical scattering can occur when one wants to know at which lattice point the electron is reflected, because then the de Broglie wavelength becomes very short on account of a large amount of momentum which the electron receives. No mention is made of the “definite experiment” to confirm the lattice point at which the electron is scattered.
§3. **Heisenberg’s interpretation of quantum-mechanical laws**

Considering that a “direct and intuitive understanding” of the commutation relation is seen in the uncertainty relation, Heisenberg now attempts to interpret quantum-theoretical laws on the basis of the uncertainty relation. “Summarizing and generalizing the results” of the “thought experiments” described above, he makes firstly the following “assertion” — “All concepts that are used to describe a mechanical system in classical theory can be exactly defined also for atomic processes analogously to the classical concepts”. This is however conditional. Namely, he says that “the experiments which serve such a definition carry purely empirically an uncertainty in them, if we demand the simultaneous determination of two canonically conjugate quantities”.

In matrix mechanics, which described the phenomenological relations between quantum-theoretical amplitudes with matrices, the usual kinematical concepts were “broken”, as Heisenberg says, so that it was not clear what meanings the matrices $q$ and $p$ had, though they were supposed to correspond to the position coordinate and momentum. On the other hand wave mechanics, which treated the de Broglie wave “in intuitionally concrete form”, gave substantialistic grounds for $q$ and $p$, but the position coordinate $q$ became the variable of the wave function $\psi(q)$ as the place of existence of the wave, making its relation to the position coordinate of the particle unclear.

Born’s theory, which takes $|\psi(q)|^2 dq$ to be the probability of finding the particle in the interval between $q$ and $q + dq$, revives the position coordinate of the particle in the quantum theory. Dirac-Jordan’s theory, in which the wave function $\psi_\eta(q)$ of the particle in a state $\eta$ is taken to be an element of the transformation matrix, or the projection $S(\eta, q)$ of the state vector, gives at the same time the probability $|\psi_\eta(p)|^2 dp$ of finding the momentum of the particle in the interval between $p$ and $p + dp$ in terms of $\psi_\eta(p) = S(\eta, p)$. $S(\eta, q)$ and $S(\eta, p)$, being projections of the state $\eta$ different from each other, show respective behaviours, but depend on each other on account of Eq. (4). Heisenberg’s attempt described above amounts to an endeavour to express “intuitively” these circumstances by means of the quoted “assertion” and the “uncertainty”.

Heisenberg considers next, by viewing the quantum theory in this way, that “it suggests itself that here the quantum theory is to be compared to the special relativity theory”. He says that in the relativity theory “the word ‘simultaneous’ cannot be defined other than by those experiments into which the
propagation velocity of the light essentially enters”, and by this “room has been created for the postulate of the constant light velocity, so that this postulate is not in contradiction with the use of the words ‘position, velocity, time’.

Heisenberg continues to say that “there is something similar about the definition of the concepts ‘position, velocity of an electron’ in the quantum theory”. He considers that “all experiments which we can use for the definition of these words contain necessarily the inaccuracy given by Eq. (1)”, and states that “this inaccuracy creates room for the validity of the relations which find their concise expression in the quantum-theoretical commutation relations”, stressing the similarity between the constancy of light velocity and the uncertainty relation.

There are certain difficulties, however, in comparing the uncertainty relation to the constancy of light, as may be seen from Heisenberg’s usage of the respective phrases, “creates room for the postulate of the constant light velocity” about the relativity theory, and “creates room for the validity of the relations” about the quantum theory. From the postulate of constant light velocity, or the principle of the constancy of light velocity, there derives the Lorentz transformation, which is at the core of the special theory of relativity. On the contrary, from the uncertainty relation, the commutation relation or the Schrödinger equation cannot be derived.\(^8\) The former is a conclusion of the latter, and is a form which the essentialistic quantum-mechanical law brings to the phenomenon. It may be said that the thought experiment of “\(\gamma\)-ray microscope” shows us such a phenomenal form well in an intuitive way, by means of the wave and particle natures of light quantum. In the derivation of Eq. (2), the particle and wave natures of atom is used too.

The comparison by Heisenberg of the uncertainty relation to the principle of the constancy of light velocity becomes the cause giving rise later to some people calling the uncertainty relation the uncertainty principle, though Heisenberg uses the word “relations” in comparison to the word “postulate”. Furthermore, the term “uncertainty principle” is brought into philosophical discussions of epistemology, being separated from concrete investigations of the quantum-mechanical laws and atomic phenomena, and is interpreted arbitrarily from the standpoints of agnosticism, skepticism, empiricism, Machism, symbolism and so on, giving rise to various kinds of confusion.\(^9\)

\(^8\) From other principles in physics, for example, the Hamilton principle, the Fermat principle, the Boltzmann principle, etc., their respective characteristic laws can be derived.

Now, Heisenberg states again, in his paper under consideration, that Eq. (2) gives a "direct explanation of the well-known relations \( Et - tE = \hbar/2\pi i \) or \( Jw - wJ = \hbar/2\pi i \), and on account of these uncertainty relations "the phases are principally not determined in a definite stationary state". According to him, an orbit of an electron is "a series of points in space which the electron takes one after the other as position" and, therefore, "for example, the often used expression ‘1s-orbit of the electron in hydrogen atom’ has no meaning from our viewpoint".\(^{10}\)

Heisenberg considers that "on the other hand the position measurement can be carried out on many atoms in the 1s-state", and that for this "there must be given such a probability function for the position of the electron in a definite state, for example 1s of the atom, that corresponds to the average value of the classical orbit over all the phases and is ascertainable by measurements as exactly as one likes". By saying that "according to Born this function is given by \( |\psi_{1s}(q)|^2 \), Heisenberg tries to explain the interpretation of the wave function as the probability amplitude.

Namely, he thinks that "also in classical theory the probability of a definite position of the electron could be given, as long as we do not know the phases of the atom". Then, he thinks that "classically the phase of the atom can always be determined by experiments", while the phases of the atom in a quantum stationary state "are principally not determined", and interprets the appearance of the probability in quantum mechanics "correspondingly" with the classical theoretical probability in the case of no knowledge of the phases.

Heisenberg’s thought that "on many atoms" the position measurement can be carried out and for this the probability must be given is similar to the way of thinking that introduces probability into classical statistical mechanics. In classical statistical mechanics, the \( q \) and \( p \) of each particle is in principle determined at any time, but is unknown and hidden because of the largeness

\(^{10}\)Heisenberg, about Schrödinger’s paper we mentioned in Sec. 5, §1, Chap. 2 of the present volume [Ref. 51], Chap. 2, in which Schrödinger compared quantum mechanics with classical mechanics with the use of a wave packet calculated by superposing eigenfunctions of a harmonic oscillator, states that he “does not believe that Schrödinger’s consideration hits the essence of the problem”, because the harmonic oscillator is a special case in which any difference between two eigenvalues of energy is an integer multiple of a constant. Heisenberg considers that “the transition from the micro- to macro-mechanics” as discussed by Schrödinger should be based on the circumstance, that in a highly excited state of an atom “the size of the orbit” is so large that \( q_1 \) could be small compared with the size of the atom, even if \( q_1 \) is taken large enough to make \( p_1 \) small.
of their total number. Consequently, average quantities are taken over a large number of particles, and the probability is introduced for the \( q \) and \( p \) of a particle to be found in a small domain in phase space.

The wave function \( \psi_1(q) \) is, however, concerned with an atom. In this connection, Heisenberg himself writes: "When we know that an atom was in the state of energy \( E_n \), we can bring this experimental fact to the expression that we assign a wave function to the atom". Thus, after having interpreted \( |\psi(q)|^2 \) in connection with many atoms, Heisenberg adds the following words — "We have not adopted that, in contrast to the classical theory, the quantum theory is an essentially statistical theory in the sense that from exactly given data only statistical conclusions could be drawn. Indeed the well-known experiments by Geiger and Bothe for example speak also against such assumptions. In all the cases in which relations exist in classical theory between two quantities, which are really all exactly measurable, there hold rather the corresponding exact relations also in quantum theory (momentum-energy law)".

Bothe-Geiger's experiments cited here are those we have mentioned in Sec. 3, §3, Chap. 5, Vol. II of the present book. Their experiments, together with Compton-Simon's experiment we have mentioned in the same section as above, confirmed (in 1925) that the law of energy-momentum conservation holds in each individual Compton effect, negating the theory of statistical conservation of energy-momentum proposed by Bohr-Kramers-Slater (in 1924). Bohr, who had had some doubt about the theory of light quantum, supposed (in 1923) that there would be "such a mechanism that the total reaction of a number of atoms is the same as that of a number of harmonic oscillators in the classical theory", and introduced further, with Kramers and Slater, "a set of virtual harmonic oscillators" to propose the theory of statistical conservation of energy-momentum (cf. §§1 and 2 of the section mentioned above).

If one interprets \( |\psi(q)|^2 \) in a statistical way only, one is not permitted to calculate the elementary process of the Compton effect with the use of \( \psi(q) \) as in Gordon's paper (in 1926), but is led to a viewpoint similar to Bohr-Kramers-Slater's. This circumstance forms the background of the words added by Heisenberg. From his paper it is not clear, however, in what logic the fact that the state of a particle is represented by \( \psi(q) \) is consistent with the fact the \( |\psi(q)|^2 \) gives the statistical probability of finding the particle.

Heisenberg further adds a statement to deny the causality on the ground of the uncertainty relation. Namely, he says that "when we know well the present, we can calculate the future", is not the final remark but the false
assumption”. Then, he says that “we cannot get to know the present in all the determinative pieces”, and “for this reason all the perception is a choice from an enormous amount of possibilities and a restriction of the future possibilities”. Heisenberg continues to say that “because the statistical character of the quantum theory is now very closely connected with the inaccuracy of all perception, one could be led to the supposition that, behind the perceived statistical world, there still hides a ‘real’ world in which the causal law is valid, but such speculations seem to us, this we stress explicitly, infertile and senseless”, denying the so-called hidden variables.

Thus, from his Machismic and phenomenological view, expressed here again that ‘physics should formally describe only the relation of perceptions”, he states that “one can rather characterize the true facts of the matter much better as follows” — “Because all the experiments are subjected to the quantum mechanical laws and therefore to Eq. (1), by quantum mechanics the invalidity of the causal law is definitely established”. That is, he considers that, because the perception of the present is in principle restricted by the uncertainty relation, it is impossible to calculate the future.\(^\text{11}\)

It should be noted, however, that even though at the time \(t = t_0\) a particle is at the position \(q = a\), its wave function being given by \(\psi(q) = \delta(q - a)\) according to Dirac and so its \(p\) being completely indeterminate, at any later time the wave function is uniquely determined by the time-dependent Schrödinger equation \(\frac{i\hbar}{\partial t} \psi = H \psi\), and from it the form of motion of the particle can be known, so that causality is never destroyed by the uncertainty relation at \(t = t_0\). We shall treat this point again in later sections. It may suffice here to note that Heisenberg’s thought described above serves as an example to show that “physics itself is utterly different from interpretations of physics”.\(^\text{12}\) In many cases, “the interpretation of physics by a physicist is restricted by his limited view of world and by his logic”.\(^\text{12}\)

---

11) In “VI. The Statistical Interpretation of the Quantum Theory” of the book “Die Physikalischen Prinzipien der Quantentheorie” [The Physical Principles of the Quantum Theory; S. Hirzel, Leipzig (1930)], in which the lectures given by Heisenberg at Chicago University in the spring of 1929 are collected, Heisenberg says that “two quantum mechanical quantities are causally connected with each other only when the tensors corresponding respectively to them have mutually parallel principal axes, · · · in all the other cases they have no causal relation”. “Tensors” mean here matrices in the Hilbert space. In this way, he supposes that the causal law holds only for the case of such physical quantities that are commutable and are therefore determinable together, denying in general causality.

12) M. Taketani, Ref. 9) of the present chapter.
§4. The uncertainty relation and Bohr’s idea of complementarity-reciprocity

Bohr attempted, in his paper "The Quantum Postulate and the Recent Development of Atomic Physics" (N. Bohr, Die Naturwiss. 16 (1928), 245) published on Apr. 13, 1928, to interpret the characteristics of quantum mechanics in terms of "complementarity" and "reciprocity". This paper covers thirteen pages, but is printed in two columns and in small type every page, so that it would correspond to a paper of thirty-three pages in Zeit. f. Phys. in which Heisenberg's paper treated in the last section was published. Its content is "essentially" identical with that of Bohr's lecture given at the Volta Celebration held in Como on Sep. 16, 1927, as he stated in a footnote. This paper is divided into seven sections, §1. Quantum Postulate and Causality, §2. Action Quantum and Kinematics, §3. Measurement in the Quantum Theory, §4. Correspondence Principle and Matrix Theory, §5. Wave Mechanics and Quantum Postulate, §6. Reality of Stationary States, §7. The Problem of the Elementary Particles. It depends much on Heisenberg's paper we treated above and may better be called an essay-like review article.

In the beginning of §1, Bohr says that "the quantum postulate results in a renunciation of the causal space-time description of atomic phenomena", and in the end of §1, that "the difficulties, which in the quantum theory stand in the way of a causal space-time description and have been the object of interest for a long time, have recently moved into the foreground of interest through the development of the new symbolic method". He then says, citing Heisenberg's paper taken in the last subsections, that "Heisenberg has particularly emphasized the peculiar reciprocal uncertainty which is attached to every measurement of atomic quantities".

"Before we enter further into his considerations, it will however be appropriate to see how the complementary feature of the description standing out in this uncertainty appears to be unavoidable, even in an analysis of the simplest concepts which lie at the base of the interpretation of experiences". Saying so, Bohr goes to his explanation of the uncertainty relation in §2 and of Heisenberg's "thought experiment" in §3.

Now, the terms "complementary" and "reciprocal", which are peculiarly used by Bohr, are brought in for the first time by him in the middle of §1 of this paper. Namely, Bohr says there the following — "On the one side, the definition of the state of a physical system demands ······ the exclusion of all

13) N. Bohr, Die Naturwiss. 16 (1928), 245.
external influences; then, however, according to the quantum postulate every possibility of observation is also excluded and, above all, the concepts of time and space lose their direct sense. If we allow, on the other hand, to make observations possible, eventual interactions with suitable external measuring means not belonging to the system, an unambiguous definition of the state of the system is not possible any longer according to the nature of things, and causality in the usual sense is out of question. According to the essence of the quantum theory, we must therefore content ourselves with regarding the space-time representation and the demand for causality, the uniting of which is characteristic of the classical theory, as complementary but each other excluding features of the description of the content of the experience that symbolize the idealization of the possibility of observation or definition”.

As we have described before, Heisenberg considered that, in order to make clear “what should be understood” by the phrase the position of an electron, “one must state definite experiments” and derived the uncertainty relation $\Delta p \Delta q \sim h$ from a thought experiment on $\gamma$-ray microscope. Borrowing Bohr’s wording, we may say that Heisenberg derived a relation for $\Delta p$ and $\Delta q$ which are exclusive of each other as well as reciprocal or complementary to each other, by considering a symbolic experiment which idealizes the possibility of definition.

Also, Heisenberg considered, on the basis of the uncertainty relation, that such an expression as the 1s orbit of an electron “has no meaning”, because an orbit of an electron is “a series of points in space which the electron takes one after the other as position”, and supposed that “the invalidity of the causal law is definitely established”, for the reason that the perception of the present is in principle restricted, on account of the impossibility of determining simultaneously a pair of canonical conjugate quantities. In Bohr’s way of expressing as seen in §3 of his paper, Heisenberg considered that “not only the measurement of the position coordinates of a particle is connected with a finite change of the dynamical variables, but also the fixing of its position means a complete break in the causal description of its dynamical behaviour, in exactly the same way as the knowledge of its momentum is always gained at the cost of an unfulfillable break in the pursuit of its space-time transmission”, and “just this circumstance brings clearly the complementary character of the quantum-theoretical description of atomic phenomena to light”.

As we have mentioned before, Heisenberg made use of the particle and wave natures of light quantum and a material particle in the derivation of the
uncertainty relation by means of his thought experiment. He called the antago­

nism between the corpuscular and undulatory theories due to this duality “the struggle of the views about particles and waves”, as we have seen pre­

viously. Also, we have made clear that quantum mechanics was established through the unified grasping of the antagonism between the dual natures, and have pointed out that the uncertainty relation is a conclusion from quantum mechanics, which can give not only the relation $\Delta p \Delta q \geq \frac{\hbar}{4\pi}$ but also the values of $\Delta p$ and $\Delta q$ themselves in definite phenomena. The uncertainty relation is, so to speak, a projective representation of the unifiedly grasped duality. As far as this means, the uncertainty relation might be an expression of “com­

plementary but to each other exclusive features” — since the product of $\Delta p$ and $\Delta q$ is limited to be at least $\frac{\hbar}{4\pi}$, just as the sum of complementary angles is limited to be $\pi/2$.

§5. “Copenhagen spirit” and its limit

Though Bohr’s idea of complementarity-reciprocity is connected with Heisen­

berg’s uncertain relation as we have seen in the last subsection, Bohr supposes the latter to be principal. Namely, he stresses in §1 of his paper that “in the case of the description of atomic phenomena the quantum postulate puts us in front of the task of developing a ‘complementarity theory’”. From this standpoint of Bohr, the antagonism between the wave and particle natures, or between the corpuscular and undulatory theories, is also “complementary”. In §5 of his paper, Bohr regards each of wave mechanics and matrix mechanics as “a symbolic paraphrase, which is appropriate to the demands in the quantum theory of the corresponding problem of motion in classical mechanics”, and says that “the two formulations of the problem of interaction may be de­

scribed, in view of their starting-points — the wave or particle conception of free individuals — as complementary”.15)

At the end of his paper, Bohr considers that, in carrying out “the task of developing a complementarity theory”, “the obstacles which we meet on the way stem, above all, from the fact that, so to speak, every word of a language is tied to the forms of intuition”, and says that “in the quantum theory this

---

14) Complementary is a derivative of complete.
15) We remember the fact that Bohr was in doubt of the theory of light quantum, and, from the standpoint of wave theory, proposed Bohr-Kramers-Slater’s theory of statistical conservation of energy-momentum. Does this mean that Bohr proposed a theory which was not opposed but complementary to the theory of light quantum?
difficulty comes up immediately to us in the question of the necessity of the feature of irrationality inherent to the quantum postulate”. Thus, Bohr concludes his paper by saying “I hope that the concept of complementarity will be suitable to characterize the existing situation, which should show a close analogy with the general difficulties for human beings in forming concepts which are based on the separation between the subject and the object”.

In this way, Bohr attempts to grasp the antagonism between the subject and the object as a case of complementarity. As is seen from his expression in §1 of his paper that “not to each other contradictory, but complementary understandings of phenomena, which only together perform a natural generalization of the classical way of description”, the subject and the object are not considered by Bohr to be in opposition and contradiction to each other, so that the boundary between the subject and the object becomes obscure in Bohr's mode of thinking. In fact, in §1 of his paper, having mentioned that “the quantum postulate means that every observation of atomic phenomena demands an interaction with the means of measurement which is not to be neglected”, Bohr states the following.

“Anyway, the concept of observation contains an arbitrariness, since it is essentially based on which objects are reckoned as the system to be observed about it. Of course in the end every observation will be able to be traced back to our sensory perceptions. The circumstance that in the interpretation of observations one must always bring in theoretical ideas, however, means that in every individual case it is a question of the appropriateness which place one introduces the concept of observation and the ‘irrational’ feature of description associated with the quantum postulate into”.

Namely, Bohr thinks that the boundary between the subject (the observer) and the object (the observed object) in an observation can be put anywhere as long as it suits the purpose. This viewpoint of Bohr is described more clearly in his paper “Action Quantum and Description of Nature”, published on Jun. 28, 1929, as follows — “On the one hand the description of our activity of thinking demands the confrontation between an objectively given content and an observing subject, whereas on the other hand, as is seen already from such a statement, no strict distinction between the object and the subject is to be given straight since even also the latter concept belongs to the content of thought”. Although his viewpoint is described here more clearly, the logic used

---

16) N. Bohr, Naturwiss. 17 (1929), 483.
is simple and formal, being similar to that in the play of words that there is no complete map because on a complete map its location would be indicated.

Bohr’s opinion that the boundary between the observed object and the observer is so arbitrary that it can be shifted up to our sensory perception has been adopted by v. Neumann as “the principle of psycho-physical parallelism”, according to which v. Neumann’s theory of observation is presented in 1932, as we shall describe in the next section. Actually, between the observed object and the observer there exists an objective boundary, as we shall discuss also in the next section.

Heisenberg wrote a preface dated Mar. 3, 1930 to his book “Die Physikalischen Prinzipien der Quantenteorie” cited before,17 in which he called Heisenberg-Bohr’s way of interpreting quantum mechanics “the Copenhagen spirit”. That is, in his preface Heisenberg emphasized as follows — “In the present book no matter is contained other than those found in previous papers, especially in Bohr’s well-known studies. I think that the aim of the present book is attained when the present book contributes something to the spread of ‘the Copenhagen spirit of the quantum theory’, if admitted to be called so. The Copenhagen spirit is indeed that which has given the guide to the whole development of the new atomic physics”.

Thus, there became prevalent an interpretation of quantum mechanics, called “the Copenhagen spirit of the quantum theory”, which in the main consists of the uncertainty relation, the negation of causality, the making much of statistical character and “the principle of psycho-physical parallelism”, and is adorned with the wording of complementarity. As to the making much of statistical character, Bohr says in his paper in 1928 cited above, that “for further pursuit of the correspondence of the quantum laws to classical mechanics the emphasis of the statistical character of quantum-theoretical description demanded by the quantum postulate has been of fundamental significance”. Also, he says in the paper in 1929 that “it cannot be surprising that all logical applications of the quantum theory have always been essentially statistical problems”, and states that “this character peculiar to the quantum theory finds expression in an impressive way in the renewed discussion about the essence of light and of the constitution of matter”.

As we have seen above, Bohr does not consider those contradictions, which exist in the antagonism between the particle and wave natures when unifiedly

---

17) Cf. Footnote 11) in §3 of the present section.
grasped, to be contradictions as they are, but interprets them in terms of complementarity. As to “the apparent contradiction which the discussion about the essence of light and the constitution of matter has revealed”, he says in the paper in 1929 that “in order to emphasize that it is not a matter of real contradictions, the word complementarity was proposed in an earlier article of the author (Naturwiss 16, 245 [1928])”. No contradiction could be made to disappear, however, no matter how rhetorically it is stated. Thus, Bohr continues to say that “in order to express the meaning of the facts of the matter in question, however, the word reciprocity should be better suited”. In this way, Bohr moves about back and forth between the complementarity and reciprocity, because he does not take antagonisms as definite contradictions.

“The physical knowledge does not proceed uniformly in the sense that ‘how it becomes then and then’, but does proceeds along the series of coiling turns of three stages”,18) namely, the phenomenological, substantialistic and essentialistic stages. In the present book we have traced in detail how quantum mechanics passed through such series of coiling turns to be established. Bohr’s complementarity comes from a view of quantum mechanics that, “being unable to grasp an object of stereo-structure as it is, so to speak, one projects it onto a plane and sees the projection in a formal way”.19)

It is because Bohr sees the developments in physical cognition and their projections onto a plane on equal footing, that he says that in quantum mechanics “we meet the obstacles which stem from the fact that every word of a language is tied to the forms of intuition”, as we have quoted above. Every time when physical cognition arrives at an essentialistic stage, one or more new essential concepts such as state vector in quantum mechanics are obtained. They are of natures different from those of the concepts, that were obtained in the previous coiling turns and have become intuitively understood. They are however not unconnected with the previously obtained concepts. Bohr’s theory is incapable of grasping this relation of stereo-structure.

In his paper in 1929 cited above, Bohr mentions also that “the conscious analysis of any concept is in an exclusive relation to its direct application”, and that “with this necessity, we are familiar to resort to a, in this sense, complementary or better reciprocal way of description”. B. Spinoza, a substantialistic

philosopher in the 17th century, pointed out the contradiction included in any concept by means of the famous expression "every determination is a negation". Bohr casts a veil of the wording of complementarity-reciprocity over the contradiction of such nature.

Heisenberg's uncertainty relation has a physical meaning in its own, as a form effected by the quantum mechanical law to phenomena. On the contrary, Bohr's complementarity-reciprocity "is neither a philosophical concept nor a natural scientific concept, and is nothing but a groundless and merely phenomenological word, which Prof. Bohr uses in his own interpretation or explanation of the quantum theory".19) "The Copenhagen spirit", which in the main consists of the uncertainty relation, the negation of causality, the making much of statistical relation, and the principle of psycho-physical parallelism, and is adorned with the wording of complementarity, is a view lacking the capability for understanding of the logic of rich content and the stereo-structure of quantum mechanics. We shall return to the logic with stereo-structure of quantum mechanics in the following sections.

As is seen from Bohr's expressions quoted above, Bohr's papers under consideration are complicated by the use of synonyms and rhetorical wordings. One scarcely meets with direct and concise expressions. For example, the fact that Bohr was doubtful of the Compton effect and proposed the theory of statistical conservation of energy-momentum, which was immediately negated by Compton-Simon's experiment, is described in Bohr's paper in 1928 as follows — "Nevertheless, the conservation of energy and momentum in the case of the interaction of radiation and matter, as it comes to light in the photoelectric effect and Compton effect, finds its logical expression just through the idea of light quantum developed by Einstein. The doubts, on the one hand, about the strict maintenance of the principle of superposition and, on the other hand, about the general validity of the laws of conservation, to which this apparent contradiction has given cause, have been refuted as is well-known in a crucial way".

In the present section, we have especially selected for the purpose of analysis Bohr's thoughts which are related directly to the Copenhagen spirit, out of a flood of words, so to speak, in Bohr's two papers. About his thought on other problems, we note here the following two additional points. First, Bohr uses in his 1929 paper the expression that "the theory of relativity, which was to reveal the subjective character of all the concepts in classical physics, just through a far-reaching deepening of the problem of observation". This shows that Bohr
considers the difference in time and length due to the difference in reference systems of coordinates to be the problem of the subject of the observer.

Second, Bohr brings his interpretation of the uncertainty relation into the “field of psychology”. Namely, he says, also in the same paper, that “we have learned that every attempt to acquire a knowledge of atomic processes will be accompanied by a principally uncontrollable intervention in its course”, and “we must therefore be prepared for that an attempt to observe brain processes would bring with itself an essential change of the accompanying feeling of the will”.

Bohr continues to say that “although here it can be a matter only of more or less relevant analogies for the moment, one will hardly be able to free himself from the convection that, in the facts inaccessible to our ordinary view which are revealed by the quantum theory, we have got in hand a way of examining general problems of human thinkings”. It should be mentioned that the introduction of an interpretation of the uncertainty relation of this kind, and still more of the projected view as we have analyzed above, into the field of psychology (brain physiology?) is just a “subjective” speculation.

2. v. Neumann’s Theory of Observation and Critique of It

§1. v. Neumann’s statistical operator and “mixed” aggregate

§1.1 Statistical Operator

v. Neumann proposed a theory of observation following the so-called Copenhagen spirit in the last chapter “VI. The Measuring Process” of his book “Mathematical Foundations of Quantum Mechanics”\(^{20}\) published in 1932. His theory of observation was developed by means of an operator which he named “the statistical operator” in this book. This operator was introduced in his paper “Probability-Theoretical Construction of Quantum Mechanics”,\(^{21}\) received on Nov. 11, 1927, to attempt “an inductive construction” of the statistical interpretation of quantum mechanics.

In this paper, v. Neumann says that the method used previously in quantum mechanics was “essentially deductive”. Namely, he says that the absolute value

squared of a certain expansion coefficient of the wave function, or also that of the wave function itself, “was quite dogmatically equated with probability” and “the agreement with the experience was afterwards verified”. This indicates the process that started from the application by Born of wave mechanics to the collision problem, which we have described in Sec. 5, Chap. 2 of the present book, and, by way of the theoretical analyses and formulations by Heisenberg, Dirac, Born, London and Jordan we also mentioned in Secs. 1 and 2, Chap. 3, ended at the grasp of the concept of probability amplitude, which were applied to various problems.

v. Neumann continues to say that “a systematic derivation of quantum mechanics from empirical facts, that is, an inductive founding, was however not given”. Thus, he considers that “in the present paper such an inductive construction is to be attempted”. In doing so, he takes Heisenberg’s statistical interpretation, which we have treated in the last section, as the base of this study. Namely, v. Neumann thinks that “the statistical, ‘acausal’ nature of quantum mechanics is caused singularly by the (principal!) insufficiency of measurement”, as he notes in the summary at the end of the paper by citing Heisenberg’s paper on the uncertainty relation.

However, different from Heisenberg, v. Neumann does not negate generally the causality from the uncertainty relation, and “notices” also in the summary that “a system leaving itself as it is (which is bothered by no measurement whatever) develops with time completely causally: if one knows its state \( \psi \) at the time \( t = t_0 \), one can calculate it for later times according to the time-dependent Schrödinger equation”.

“About experiments the statistical character is not to be avoided: for every experiment there exists, indeed, a state suited to it, in which its result is definite; but for every state there exist unsuited experiments, the carrying out of which wrecks the same — and according to laws of chance a suitable state is generated — .” v. Neumann’s words are to mean that if a physical quantity is measured about the system in an eigenstate of the quantity the result is definite, whereas if it is measured about the system in an arbitrary state the system transits into an eigenstate of the same quantity, and it can be answered only in terms of the probability which eigenstate the system transits to.\(^\text{22}\)

\(^\text{22}\)In relation to this problem, v. Neumann discussed in III.3. of his book cited above Compton-Simon’s experiment as follows — In this experiment the energy and momentum of a light quantum as well as an electron are measured, but the measurement \( (M_1) \) for detecting a light quantum and the measurement \( (M_2) \) for detecting an electron are carried out separately. The result is independent of the order \( M_1 \) and \( M_2 \). In principle, three stages can be considered about the causality or the acausality. Firstly, the measured quantity \( (\mathcal{R}) \) is
v. Neumann's Theory of Observation and Critique of It

In this way, v. Neumann sees "the principal insufficiency of the measurement" as the cause of the statistical character of quantum mechanics, and takes the statistics of measurements as the starting-point to treat "inductively" quantum mechanics. But, the statistical arrangement of the results of measurement is no more than a simple application of statistics, and the fundamental laws of quantum mechanics cannot be derived, as we have described in the last section, from the uncertainty relation. It is v. Neumann's attempt to present a formulation of the statistics of measurement, and to relate it to the "deductive" laws of quantum mechanics. The "statistical operator" is introduced for this purpose.

Now, in order to treat the statistics of the results of measurement, v. Neumann considers an aggregate \( \{S_1, S_2, \ldots\} \), which has as its specimens "very (about infinitely) many" systems, each of which has the same "physical structure" as the system \( S \). The state of every system in this aggregate is "arbitrary". He considers further an aggregate \( \{S'_1, S'_2, \ldots\} \) consisting of systems which are of the same physical structure as \( S \). Of any system \( S' \) in this aggregate, it is assumed that "one knows something". For example, it is assumed to be known that "the value of the quantity \( a \) lies in the interval \( I \)". In this case, \( S'_1, S'_2, \ldots \) mean "\( S_1, S_2, \ldots \) subjected to the experiment of 'measurement of \( a \)'".

"According to the fundamental principles of the probability theory, one gains the same statistics if one examines, instead of \( \{S_1, S_2, \ldots\} \), its (arbitrarily selected) subsystem, which itself has just many elements, but in comparison with \( \{S_1, S_2, \ldots\} \) is as small as one likes". This concerns the statistical accuracy of the experiment. He says, in IV.1. of his book cited above, that "this follows from the so-called law of large numbers, that is, the Bernouilly theorem".

v. Neumann continues to say, in the paper under consideration, that "every 'knowledge' of \( S' \), that is, every aggregate \( \{S'_1, S'_2, \ldots\} \), corresponds to an assignment of an expectation value to every quantity \( a \) (that is its mean value occurring in \( \{S'_1, S'_2, \ldots\} \))". "The 'knowledge' of \( S' \), which is the statistical composition of \( \{S'_1, S'_2, \ldots\} \), is completely described, if the assignment

\[
a \leftrightarrow \text{Expectation value of } a = E(a)
\]

entirely statistical so that the result of the first measurement has no relation whatever to that of the next measurement. Secondly, the result of the first measurement is statistical, but the result of the next measurement is the same as that of the first. Thirdly, \( R \) is from beginning causally determined. "From the first basic idea a statistical theory of elementary phenomena was built, which Bohr, Kramers and Slater worked out. Compton-Simon's experiment can be seen as a refutation of this view."
is given for all quantities \( a \)”. Now, v. Neumann considers that “for the calculation it is naturally suitable to assign a normal operator to the expectation value of the quantity belonging to the operator”. Namely, he supposes that the mean value of \( a \), \( \mathbb{E}(a) \), occurring in \( \{S'_1, S'_2, \ldots\} \) is equal to the expectation value of an operator \( R \), \( \mathbb{E}(R) \). But the expectation value of an operator is premised here only of its existence, being defined in the following way.

In relation to the abstract Hilbert space spanned by a complete set of orthonormal functions, the matrix elements of \( R \) are denoted by \( r_{\mu\nu}'s \). v. Neumann introduced the operators \( A^{(\kappa)}, B^{(\kappa\lambda)} \) and \( C^{(\kappa\lambda)} (\kappa < \lambda) \), whose elements are given by

\[
A^{(\kappa)}_{\mu\nu} = \begin{cases} 1 & (\mu = \nu = \kappa) \\ 0 & \text{(otherwise)} \end{cases}, \\
B^{(\kappa\lambda)}_{\mu\nu} = \begin{cases} 1 & (\mu = \kappa, \nu = \lambda) \\ 0 & \text{(otherwise)} \end{cases}, \\
C^{(\kappa\lambda)}_{\mu\nu} = \begin{cases} i & (\mu = \kappa, \nu = \lambda) \\ -i & (\mu = \lambda, \nu = \kappa) \\ 0 & \text{(otherwise)} \end{cases}.
\]  

\( R \) can be expressed in terms of \( A^{(\kappa)}, B^{(\kappa\lambda)} \) and \( C^{(\kappa\lambda)} \) through the relation

\[
R = \sum_{\mu} r_{\mu\mu} A^{(\mu)} + \sum_{\mu < \nu} \text{Re}(r_{\mu\nu}) B^{(\mu\nu)} + \sum_{\mu < \nu} \text{Im}(r_{\mu\nu}) C^{(\mu\nu)}.
\]  

(13)

If one “needs” the expectation value of the sum of operators to equal the sum of the expectation values of each operator, one has, by introducing the operator \( U \) whose elements are given by

\[
\begin{align*}
  u_{\mu\mu} &= \mathbb{E}(A^{(\mu)}) \\
  u_{\mu\nu} &= \{ \mathbb{E}(B^{(\mu\nu)}) + i\mathbb{E}(C^{(\mu\nu)}) \}/2 \quad (\mu < \nu), \\
  u_{\nu\mu} &= \{ \mathbb{E}(B^{(\mu\nu)}) - i\mathbb{E}(C^{(\mu\nu)}) \}/2 \quad (\mu < \nu)
\end{align*}
\]  

(14)

the relation

\[
\mathbb{E}(R) = \sum_{\mu\nu} u_{\nu\mu} r_{\mu\nu} = \text{Trace}(UR).
\]  

(15)

v. Neumann calls this operator \( U \) the statistical operator in his book cited above.
§1.2 "Uniform" aggregate and "mixed" aggregate

v. Neumann calls an aggregate \{S_1, S_2, \ldots\} a "pure" or "uniform" aggregate if all the systems in it are in one and the same state. When a uniform aggregate is divided arbitrarily into two sub-aggregates,\(^{23}\) the expectation values of \(R\) in the sub-systems \(E'(R)\) and \(E''(R)\) are to be equal to each other, so that

\[
E(R) = E'(R) = E''(R).
\]  

(16)

In this case, \(U\) is given by those of the sub-aggregates \(U'\) and \(U''\) to be

\[
U = U' + U''.
\]  

(17)

From this relation there should result the condition that \(U' = c'U, U'' = c''U\) ["From the standpoint of relative probability and relative expectation the values \(U\) and \(cU\) \((c > 0)\) do not essentially differ".]

v. Neumann notes that there exists a function \(f_0\) for which \(U f_0 \neq 0\) on account of the relation \(U \neq 0\), and takes the function \(\varphi\) to be the \(U f_0\) normalized to unity. Defining the projection operator \(P_\varphi\) by\(^{24}\)

\[
P_\varphi f = (\varphi, f)\varphi, \tag{18}
\]

he shows that the condition mentioned above is satisfied only when\(^{25}\)

\[
U = P_\varphi. \tag{19}
\]

In this case, with the use of such a complete set of orthonormal functions that has \(\varphi\) as one of its members, \(\varphi(= \varphi_1), \varphi_2, \varphi_3, \ldots, E(R)\) is calculated to be

\[
E(R) = (\varphi, R\varphi). \tag{20}
\]

Here, \((\varphi, R\varphi) = \int \varphi R\varphi d\tau\) is the quantity that is taken to be the expectation value of \(R\) in the method of quantum mechanics referred to by v. Neumann as "deductive". All the systems \(S'\)'s are therefore known to be in the state represented by the wave function \(\varphi\). Thus the meaning of \(E(R)\) for a uniform aggregate is obtained. However, \(\varphi\) cannot of course be specified only by the above mentioned condition which he sets up.

\(^{23}\)Arbitrarily as long as every sub-aggregate obeys the law of large numbers.

\(^{24}\)\((\varphi, f)\) denotes the inner product of \(\varphi\) and \(f\), \((\varphi, f) = \int \varphi f d\tau\).

\(^{25}\)v. Neumann uses the fact that \(U'\) defined by \(U'f = (U f_0,f)/(f_0,U f_0)U f\) satisfies this condition. The same proof as is given in his paper under consideration is seen in his book cited above (IV.2.).
In the next place, v. Neumann considers that if $U$ is given by

$$U = P\varphi_1 + P\varphi_2 + \cdots,$$  

(21)

"this aggregate can be gained by the mixture of the uniform aggregates of states $\varphi_1, \varphi_2, \ldots$ in the ratio of $1 : 1 : \cdots". In IV.1. and 2. of his book cited above, he considers more generally an aggregate which contains the uniform aggregates of states $\varphi_1, \varphi_2, \ldots$ in the ratio of $w_1 : w_2 : \cdots$ so that

$$U = \sum_n w_n P\varphi_n,$$  

(22)

and calls this aggregate a "mixture". He also calls a uniform aggregate "a state (and indeed $U = P\varphi$ is a state)". For a mixture one has

$$E(R) = \sum_n w_n (\varphi_n, R\varphi_n).$$  

(23)

Now according to quantum mechanics, when the quantity $R$ is measured in the state represented by a wave function $\varphi$, the system transits acausally into one of the eigenstates (the eigenfunctions being $\varphi_1, \varphi_2, \ldots$). In the same book (V.1.; VI.1 and 3.), v. Neumann says, about the probability of a transition to the state $\varphi_n$, that "quantum mechanics demands $w_n = |(\varphi_n, \varphi)|^2$", and concludes from this that by the measurement the statistical operator $U$ changes into

$$U' = \sum_n w_n P\varphi_n, \quad w_n = |(\varphi_n, \varphi)|^2.$$  

(24)

On the other hand, the change in the state of the system obeys the Schrödinger equation and is given by $\varphi' = \exp(-iHt/\hbar)\varphi$, where $H$ is the Hamiltonian of the system. He concludes therefore that in this causal change $U = P\varphi$ changes into $U' = P\varphi'$, that is

$$U' = \exp(-iHt/\hbar)U \exp(iHt/\hbar).$$  

(25)

v. Neumann presents a theory of observation with the use of $U'$ given in Eqs. (24) and (25), as we shall see in the next subsection, §2. The relation $w_n = |(\varphi_n, \varphi)|^2$ in Eq. (24) is also taken according to the "deductive" method

\[\text{In his paper "Thermodynamics of Quantum Mechanical Aggregates" received on the same day as the paper under consideration, v. Neumann uses $U$ that corresponds to the supposition of $w_n \propto \exp(-E_n/kT)$ to express entropy as $S = -Nk \text{Trace}(U \ln U)$, $E_n$ being the energy of the state $\varphi_n$.}\]
of quantum mechanics, and it could not be derived by means of his "inductive" method.

Following Heisenberg's way of thinking, v. Neumann introduces the statistical aggregate of the result of measurements, and takes as the basis of his theory the fact that the statistical operator $U$, which gives the expectation value of an operator $R$ for a statistical aggregate as $E(R) = \text{Trace}(UR)$, becomes $U = P\varphi$ for a "pure" aggregate called a "uniform" aggregate, leading to the relation $E(R) = (\varphi, R\varphi)$ given in Eq. (20). On this basis he considers a "mixed" aggregate, for which $U'$ is given by Eq. (24). The characteristic and limit of v. Neumann's theory consist in that the statistics of the result of a large number of measurements is represented by an operator named the statistical operator.

§2. v. Neumann's theory of observation

v. Neumann presents, in Chap. VI of his book under consideration, a theory of observation, thinking that "the boundary between an observed system and the observer is to a great extent arbitrary", from the standpoint of "the principle of psycho-physical parallelism". He says, citing Bohr's paper [27] which we treated in Sec. 1, §5 of the present chapter, that "Bohr pointed out for the first time that the duality in the description of nature made unavoidable in formal respect by quantum mechanics is justified also as a matter of fact, and indicated the connection with the principle of psycho-physical parallelism".

According to v. Neumann, "it is a fundamental demand of the natural scientific world-view, namely the so-called psycho-physical parallelism, that it must be possible to describe the extra-physical process of the subjective apperception as if it actually occurred in the physical world". Taking the measurement of temperature as "a simple example", he considers that the boundary between the observed object and the observer in this measurement can be set in various ways: it may be the mercury container of the thermometer, the scale of the mercury thread, the retina receiving the light from the scale, or the chemical reactions in the brain cells. v. Neumann says that "the observer by no means needs to be identified with the body of the actual observer — in the above example we once even added on the thermometer, whereas at another time his eyes and nerve tracks were not added on", and that "it is the content of the principle of psycho-physical parallelism that the boundary can be moved arbitrarily deep into the interior of the body of the actual observer".

[27] N. Bohr, Naturwiss. 17 (1929), 483 [same as Ref. 16 in the present chapter].
Now, because in quantum mechanics there exist the causal change with time of a free system and the acausal change in a measurement, v. Neumann thinks that "there is the danger that the principle of psycho-physical parallelism is violated", and discusses whether the boundary between the observed system and the observer is arbitrarily movable according to quantum mechanics. For this purpose, he "divides the world into three parts I, II and III, I being the original observed system, II the measuring instrument and III the original observer". "It is to be shown that the boundary can be drawn between I and II+III as well as between I+II and III".

However, v. Neumann adds to this statement a parenthesis, in which he mentions about the above example the following cases: I = the observed system, II = the thermometer, III = the light + the observer; I = the observed system + the thermometer, II = the light + the eye of the observer, III = the observer, from the retina on; I = all up to the retina of the observer, II = his retina, nerve tracks and brain, III = his abstract "ego". Thus, he divides the world into I, II and III, by dividing "the original observer" in detail up to "his abstract ego", which is supposed to "apperceive the chemical changes in his brain cells", from the viewpoint of the principle of psycho-physical parallelism.

Anyway, his "real task" is "the proof that the same statements arise", in both the cases of the application of the acausal change in a measurement to the interaction between I and II+III and to that between I+II and III.

Let $A$ be the quantity of $I$ to be measured, and $\varphi_1(q), \varphi_2(q), \ldots$ be its normalized eigenfunctions, $q$ denoting representatively all the coordinates of $I$. If $I$ is in the state $\varphi(q)$ and the boundary is drawn between $I$ and II+III, so that $I$ is taken as "the observed system" and II+III as "the observer", the measurement of $A$ causes $I$ to transit from the state $\varphi$ to some state $\varphi_n$ with the probability $|\langle \varphi_n, \varphi \rangle|^2$.

v. Neumann says that, in the case of the boundary being drawn between I+II and III, "II is a measuring instrument which informs us the value of $A$ (in I), and the position of the pointer on a scale is a physical quantity $B$ (in II) which is actually observed by III (if II lies already in the interior of the body of the observer, there would enter, instead of the scale and position pointer, the corresponding physiological concepts, for example the retina and the image on it, etc.)".

The state of I+II is $\Phi(q, r) = \varphi(q)\xi(r)$ if $I$ is in "the unknown state" $\varphi(q)$ and II in "the known state" $\xi(r)$, $r$ denoting representatively all the coordinates of II. "The measurement (as long as it is carried out by II on I) is carried out
through an energy operator $H$ (in I+II) at time $t$. The change in the state of I+II is therefore given by

$$\Phi(q, r) = \varphi(q) \xi(r) \rightarrow \Phi'(q, r) = \exp \left( -\frac{iHt}{\hbar} \right) \Phi(q, r). \quad (26)$$

On the other hand, v. Neumann says that "seen from the standpoint of the observer, a measurement can come into question only when the following is the case". That is, if $A$ has the values $a_1, a_2, \ldots$; $B$ the values $b_1, b_2, \ldots$; and they are so numbered that $a_n$ corresponds to $b_n$, then the pair of values $a_m$ and $b_n$ will have probability 0 for $m \neq n$, and probability $|\langle \varphi_m, \varphi \rangle|^2$ for $m = n$. Saying that this is "in principle" the case, v. Neumann treats the question with a procedure outlined in the following.

Let the statistical operators for I, II and I+II be $U^I$, $U^{II}$ and $U$ respectively, and some complete sets of orthonormal functions in I and II be $\psi_1, \psi_2, \ldots$ and $\eta_1, \eta_2, \ldots$, respectively. Let $u^I_{m,m'}, u^{II}_{n,n'}$ and $u_{mn,m'n'}$ denote the respective matrix elements of $U^I$, $U^{II}$ and $U$, given in terms of $\psi_m(q)$ and $\eta_n(r)$. If $A$ has the matrix element $a_{m,m'}$ in I, it has the matrix element $a_{mn,m'n'} = a_{m,m'} \delta_{n,n'}$ in I+II. Now, v. Neumann premised that the expectation value of $A$ in the aggregate of I's is equal to the expectation value of $A$ in the aggregate of (I+II)'s.\(^{28}\) On this premise and a similar one for $B$, he gets, with the use of Eq. (15),

$$u^I_{m,m'} = \sum_n u_{mn,m'n}, \quad u^{II}_{n,n'} = \sum_m u_{mn,m'n}. \quad (27)$$

Next, expanding $\Phi'$ as

$$\Phi'(q, r) = \sum_{mn} f_{mn} \psi_m(q) \eta_n(r), \quad (28)$$

one gets from $U = P_{\Phi'}$

$$u_{mn,m'n'} = f_{mn} f_{m'n'}. \quad (29)$$

\(^{28}\)In v. Neumann's words: "The I-quantity has on the basis of a calculation in I the expectation value $\sum_{m,m'} u^I_{m,m'} a_{m',m}$, whereas the calculation in I+II results in $\sum_{m,m'} \langle \sum_n u_{mn,m'n} \rangle a_{m',m}$. In order that the two expressions are equal, it is necessary that $u^I_{m,m'} = \sum_n u_{mn,m'n}$. However, it cannot be said in general that "the two expressions are equal". v. Neumann's words mean therefore that he assumes implicitly the equality of the two expressions.
\( \bar{f}_{mn} \) being the complex conjugate of \( f_{mn} \). Therefore, by defining the operators \( F \) and \( F^* \) to be those which have respectively the matrix elements \( \bar{f}_{mn} \) and \( f_{mn} \), one has from Eq. (27) \( U^I = F^*F \) and \( U^{II} = FF^* \). \( U^I \) and \( U^{II} \) are also expressed as

\[
U^I = \sum_m w'_m \rho_m, \quad U^{II} = \sum_n w''_n \rho_n \tag{30}
\]

if they are the respective mixtures of the eigenstates of \( A \) and \( B \). Because of \( U^I \rho_m = w'_m \rho_m \) one gets

\[
U^{II}(F \rho_m) = w'_m (F \rho_m), \quad (F \rho_m, F \rho_{m'}) = w'_m \delta_{m,m'}. \tag{31}
\]

By comparing these with \( U^{II} \xi_n = w''_n \xi_n \), one sees that \( F \rho_m / \sqrt{w'_m} \) is equal to one of the \( \xi_n \)'s. By numbering the \( \xi_n \)'s in \( U^{II} \) such that \( F \rho_{\mu_k} / \sqrt{w'_{\mu_k}} \) corresponds to \( \xi_{\nu_k} \) with the same eigenvalue \( w''_{\mu_k} = w''_{\nu_k} = w_k (k = 1, 2, \ldots) \), one has

\[
F \rho_m = 0 \quad (m \neq \mu_1, \mu_2, \ldots). \tag{32}
\]

Because \( \psi_m \) and \( \eta_n \) in Eq. (28) are arbitrary, one can take \( \phi_m \) for \( \psi_m \) and \( \xi_n \) for \( \eta_n \). Thus, one gets

\[
f_{mn} = \begin{cases} \sqrt{w_k} & (m = \mu_k, n = \nu_k) \\ 0 & \text{(otherwise)} \end{cases} \tag{33}
\]

and accordingly

\[
\Phi'(q, r) = \sum_k \sqrt{w_k} \phi_{\mu_k} (q) \xi_{\nu_k} (r). \tag{34}
\]

Namely, though I and II are mixtures when the state \( \Phi'(q, r) \) of I+II cannot be represented by a single product \( \phi_m (q) \xi_n (r) \), \( \Phi' \) obtained here "brings about a one-to-one assignment between the possible values of certain quantities in I and in II". This means that from the standpoint of III "it is enough to 'look at' II". In this way v. Neumann, starting from "mixtures" of states, arrives at Eq. (34) which appears to be a superposition of states. He comes to this result because he has assumed Eq. (27), which indicates that the expectation value of \( A \) in an aggregate of I's is equal to the expectation value of \( A \) in an aggregate of (I+II)'s, and Eq. (30) which indicates that the aggregate of I's is a mixture of eigenstates of \( A \) (and similarly about II and \( B \)).

---

29) On this point the reader is further referred to Sec. 3, §4, i) of the present chapter (p. 292).
With this result, v. Neumann thinks that the observation of “the position of the pointer” in II can by itself be the measurement of a physical quantity of I. However, from the method described above, it does not follow that \( w_k \) in Eq. (34) is equal to \( |(\varphi_{\mu_k}, \varphi)|^2 \). For this reason, he puts \( \Phi \) in Eq. (26) as \( \Phi(q, r) = \varphi(q)\xi_0(r) \) and \( \Phi'(q, r) \) as
\[
\Phi'(q, r) = \Delta \varphi(q)\xi_0(r), \quad \Delta = \exp(-iHt/\hbar),
\]
and assumes that the operator \( \Delta \) can be given in the form\(^{30}\)
\[
\Delta \sum_{m,n=-\infty}^{\infty} \chi_{mn}\varphi_m(q)\xi_n(r) = \sum_{m,n=-\infty}^{\infty} \chi_{mn}\varphi_m(q)\xi_{m+n}(r).
\]
Developing \( \varphi(q) \) as
\[
\varphi(q) = \sum_{m=-\infty}^{\infty} (\varphi_m, \varphi)\varphi_m(q),
\]
he gets, with the use of Eqs. (35) and (36),
\[
\Phi'(q, r) = \sum_{m=-\infty}^{\infty} (\varphi_m, \varphi)\varphi_m(q)\xi_m(r).
\]

v. Neumann states that “our aim is thus achieved”. But, he mentions also that “about this only things of principle matter to us”, and “an abstract \( \Delta \) is given that satisfies exactly our conditions”. He is “not occupied, on the other hand, with the further question whether the \( \Delta \)'s of the simplest intuitive experimental arrangements have also this characteristics”,\(^{31}\) so that “only the existence of any such \( \Delta \) matter” to him.\(^{32}\)

If v. Neumann’s result is really one of principle, his discussions described above with the use of statistical operators are of no meaning, because Eq. (38)

\(^{30}\) About the values 0, ±1, ±2, . . . taken by \( m \) and \( n \) instead of 1, 2, . . ., v. Neumann mentions that “this has merely technical reason and is in principle indifferent”.

\(^{31}\) V. Neumann considers some experimental arrangements in his discussions of the uncertainty relation (in III.4. of his book under consideration).

\(^{32}\) As a concrete example, v. Neumann treats the case of two particles moving one-dimensionally, in which the kinetic energies are neglected by reason of their assumed large masses and the interaction energy is assumed proportional to the product of the position coordinate \( q \) of particle I by the momentum \(-i\hbar \partial/\partial r\) of particle II. With the use of \( H \) itself instead of \( \Delta \), that is, by applying the method of first order perturbation, he gives a result similar to Eq. (38). However, it should be said that the particle II treated by him is too simple to be considered a measuring instrument.
means that from the time-dependent Schrödinger equation an expression for the wave function of I+II is derived, which is of the same form as Eq. (34) and, moreover, with the necessary value of $w_k$, so that a quantity of II is seen to be observed with the probability $|\langle \varphi_m, \varphi \rangle|^2$ in one-to-one correspondence with a quantity of I. $\Delta$ in Eq. (36) should be said to belong to a very special and formal mathematical model for the purpose of getting the desired result.\textsuperscript{33)

The core of v. Neumann's theory of observation consists, rather, in the fact that discussions based on the principle of psycho-physical parallelism are made of the idea that, even if I+II is observed by III, there exists a quantity of II corresponding in a one-to-one way to the quantity of I to be observed, so that the latter quantity can be measured by means of a measurement of the former quantity. The main point of v. Neumann's theory is the assertion that, when a quantum-theoretical "composite system" is constructed from "the observed system" I and "the measuring instrument" II, any eigenvalue of the quantity to be measured of I can be known from the corresponding eigenvalue of a quantity of II indicated by II, through the statistics of the results of the measurement of "the composite system" made by "the observer" III. The Copenhagen spirit and v. Neumann's theory of observation indeed include problems "of principle", as we shall describe in the following subsections, §3 and §4, and become criticized by Schrödinger and Taketani at about the same time independently from their respective viewpoints.

§3. Critique by Schrödinger

§3.1 The case of a cat

Being unsatisfied with the interpretation of quantum mechanics based on the Copenhagen spirit, Schrödinger makes various discussions in a long paper entitled "The Present-day Situation of Quantum Mechanics"\textsuperscript{34)} to find a different interpretation. As we have described in Sec. 3, §3, Chap. 2 of the present volume, in 1926 Schrödinger criticized matrix mechanics, saying that it was of Machism and saw his own wave mechanics as the opposite to it. In our terminology, matrix mechanics was a phenomenological theory whereas wave mechanics a substantialistic theory. They were unified and quantum mechanics was established as an essentialistic theory.

\textsuperscript{33)Because $\Delta$ in the left hand of Eq. (36) contains $t$, the right hand of the same equation should also contain $t$. But v. Neumann makes no mention of the time dependence of $\hat{x}_{mn}$.}

\textsuperscript{34)E. Schrödinger, Naturwiss. 48 (1935), 807, 826, 844 [18 pages in total; cf. the explanation of Bohr's paper in Sec. 1, §4 of the present chapter (p. 248)].}
Heisenberg's interpretation of quantum mechanics was, as we have seen in Sec. 1 of the present chapter, based on a phenomenological philosophy. He attached importance to the uncertainty relation which is a phenomenal form of quantum mechanical laws, and considered that the orbit of an electron could be determined experimentally in relation to many atoms. He thought that on account of this probability was necessitated, v. Neumann formulated this idea with the use of the statistical operator, and further attempted to construct, by means of the statistical operator, a theory of observation based on Bohr's supposition that the boundary between an observed object and the observer could be moved arbitrarily.

On the contrary, Schrödinger attaches importance to the predicting ability of "ψ function" as a theoretical concept, in his paper under consideration, by thinking that in quantum mechanics "all predictions refer as before to determining items of a classical model, positions and velocities of mass points, energies and momenta, etc.", and "it is merely nonclassical that only probabilities are predicted". About the fact that the energy and angular momentum of an oscillator take only determined discrete values, Schrödinger says that "this precise statement is much more important than the knowledge of which of these values or what probability for each of them is actually present in particular cases".

Heisenberg attempted, as we have described in the last section, to get an "intuitive interpretation of quantum mechanics". Schrödinger says that "according to the wording all statements refer to intuitive models", and then "but valuable statements are scarcely intuitive about them and their intuitive features are of small value". We have pointed out in §5 of the last section (p. 253), in regard to Bohr's paper on complementarity, that every time when the physical cognition arrives at an essentialistic stage, one or more new essential concepts are obtained which are strange to, but unseparated from, the concepts obtained in the previous stages and afterward accustomed to be intuitively grasped. We may say that Schrödinger's statement quoted just above is an intuitive one, reflecting this stereo-structure of the development of physical cognition, though not based on its analysis.

About the statistical interpretation of quantum mechanics, Schrödinger says that "one brings in a so-called Gibbs ensemble" and "the body is to behave as a state picked out arbitrarily from this aggregate". He thinks that "it is natural that one attempts also to relate the always unclearcut statement of quantum mechanics to an ideal aggregate of states", but "this does not work".
Taking the example of angular momentum "as one of many", he grounds his argument on it. Namely, Schrödinger considers that "one can well choose" an aggregate of states corresponding to various admissible values of angular momentum, but this concerns a determined point. He says that for a different arbitrary point there would appear inadmissible values and "the bringing in of the aggregate therefore does not help to step further". His this argument may be said however too formal.

About "the wave function, or the $\psi$ function, that is the state vector", he says that "its law of motion, i.e., the law of its change with time, \ldots is not one just left behind the equations of motion of classical models in clarity and decisiveness", and treats the $\alpha$-decay of a nucleus. He thinks that, though the state of the nucleus is "presumably faded to such an extent and in such a way that neither the moment of the disintegration nor the direction in which the $\alpha$-particle leaves the nucleus is known", "this fadedness does not disturb us in the region of the atomic nucleus".

"The out-coming particle is described, if one wants to interpret intuitively, as a spherical wave, which emanates in all directions and incessantly from the nucleus". "The screen does not show however a constant weak area shining, but flashes at a moment at a place". Schrödinger indicates thus the wave nature of a matter particle. "Or, to tell the truth, it flashes now here, now there, because it is impossible to carry out the experiment with only a single radioactive atom". This indication is important. It points out that an $\alpha$-particle is detected now here and now there on the screen, on account of the actual experimental situation, and the $\psi$ function of the $\alpha$-particle does not concern with the statistics of many $\alpha$-particles, but with a single $\alpha$-particle.

Saying that "one can also construct entirely burlesque cases", Schrödinger proposes the famous "case of a cat". One encloses a very small quantity of radioactive substance in a Geiger counter, which is arranged to move a small hammer by means of a relay when a nucleus disintegrates to emit an $\alpha$-particle. The hammer will break a small bottle containing prussic acid. One puts all these things together with a cat into a closed iron box. The radioactivity of the substance is supposed to be so weak that perhaps one nucleus disintegrates in an hour.

One leaves the whole system untouched for an hour. If any nucleus has not disintegrated during that time, the cat is still alive. If even one nucleus has disintegrated during that time, the cat has been "poisoned". Schrödinger says that "the $\psi$ function of the whole system would therefore express that in it the living cat and the dead cat are mixed or spread in equal parts".
This case of a cat presented by Schrödinger indicates a severe critique of v. Neumann’s theory of observation, pointing out its contradiction.\(^\text{35}\) The cat corresponds to v. Neumann’s “pointer of a measuring instrument”. v. Neumann supposed that even the “positions of the pointer” are quantum mechanical eigenvalues. The life and death of the cat as “positions of the pointer” make the one-to-one correspondence to the non-emission and emission of an α-particle from the nucleus as the “measured system”, when statistics are taken over many results of measurement, and the wave function of the whole system is expressed, according to Eq. (34) in §2 of the present section, as Schrödinger says. Actually, once the cat is observed after one hour, it is known whether an α-particle is emitted or not during the time. Schrödinger states that “it is typical of these cases that an uncertainty restricted originally to the atomic area is converted to uncertainties of a gross sense, which can then be decided by direct observations”.

§3.2 The \(\psi\) function as “the catalog of expectation”

After having presented the case of a cat to indicate a severe critique of v. Neumann’s theory of observation, Schrödinger treats the problem of observation in the sections entitled “Theory of measurement, Part I”, “The \(\psi\) function as the description of state” and “Theory of measurement, Part II”, of his paper under consideration. He bases his treatment on the idea that the \(\psi\) function is a “catalog of expectation”.

Schrödinger says that “the \(\psi\) function is now an instrument for prediction of the probability of measured values” and that “the sum achieved every time of theoretically well-founded expectations of the future is embodied in it as if set down in a catalog”. As to “the catalog of expectation” there is the question of “maximality” or “completeness”. Premising the maximality of the catalog of expectation, Schrödinger derives “Proposition: With the same \(\psi\) function the system is in the same state”. “One who still thinks any difference possible with the same catalog of expectation would allow that it gives no information about all the legitimate questions”, so he proves this proposition.\(^\text{36}\)

\(^{35}\)In Schrödinger’s paper under consideration, the papers and books of Heisenberg, Bohr and v. Neumann which we mentioned in the present chapter are not altogether cited inclusive of their names.

\(^{36}\)Many wordings and modes of expression taken from the formal logic are used in this long paper of Schrödinger. In the present subsection §3, we have selected from his many discussions only those points that are directly related to the problem of observation.
About the problem of observation, Schrödinger says that "we must attempt to grasp objectively the interaction between the measured object and the measuring instrument". He considers thus two separate bodies, each of which possesses "a complete catalog of expectation — a maximal sum of knowledge — a $\psi$ function". The two bodies, when combined, "naturally possess" also a complete catalog of expectation. However, "the inverse is not true". Schrödinger says that "maximal knowledge of a total system does not necessarily include maximal knowledge of all its parts".

His argument for this statement is as follows. It is possible that when a particular measurement about the first system has this result, to a particular measurement about the second this or that statistics of expectation apply, but when the measurement in question about the first system has that result, to the measurement in question about the second a certain other expectation applies. "If such conditional propositions are found in the total catalog, the catalog cannot at all be maximal with regard to the respective systems". For this Schrödinger says that "because the content of two maximal respective catalogs is by itself already sufficient to be a maximal total catalog, the conditional propositions do not come up any more".

Now, Schrödinger supposes that two separate systems, each being "maximally known", come into a situation of mutual interaction. "The joint catalog of expectation consists initially of a logical sum of the single catalogs; during the process it develops inevitably according to known law. The knowledge remains maximal, but at the end, when the bodies have separated again, it has not split again into a logical sum of knowledge about the single bodies. That which is still to be given about it, can, very much possibly, have become less than maximal".

Schrödinger calls this "disjunctive splitting", and applies it to the case of a measuring instrument and a measured object. "From this combination of knowledge the object can be released again only through the fact that the living subject gets knowledge of the result of the measurement". "By this inspection which decides the disjunction, some discontinuity, or suddenness, happens". "However, it would not be quite right that the $\psi$ function of the object which otherwise changes according to a partial differential equation, independently of the observer, now changes suddenly owing to a mental act". "In reality, important events lie in-between, namely the mutual interaction of the two bodies, during which the object would have no private catalog of expectation and also would not be entitled to this, because it is not independent".
In this way Schrödinger stresses, contrary to the interpretation of quantum mechanics given by Heisenberg, Bohr and v. Neumann on the basis of the Copenhagen spirit, that the appearance of probability in measurement results from the objective interaction between the measured object and the measuring instrument. This should be said to be an excellent idea. However, his argument, with "the maximality of the catalog of expectation", for the occurrence of "the distinctive splitting" of the $\psi$ function as the catalog of expectation, is restricted by the limitation of formal logic. It may rather be said that "the catalog of expectation" and its "maximality" are conceived by him to stress "the disjunctive splitting" in the framework of formal logic. Schrödinger's theory is a substantialistic theory, in opposition to the phenomenological theory of observation based on the Copenhagen spirit.

This difficulty encountered in Schrödinger's theory comes to be solved when Taketani elucidates, at about the same time as Schrödinger, the logic with stereo-structure which is characteristic of quantum mechanics, as is described in the following subsection §4. Schrödinger does not treat the essential difference between the coupling of two bodies and the coupling of an observed object and the measuring instrument. This point comes to be clarified also by Taketani in 1942.

§4. Critique by Taketani

§4.1 The stereo-structural logic of quantum mechanics

Quantum mechanics is the unification, as an essentialistic theory, of the phenomenological matrix mechanics based on the corpuscular theory and the substantialistic wave mechanics based on the undulatory theory, so that it has a logical structure characteristic of this unification. In his paper "Dialectic of Nature — On Quantum Mechanics —", received on Jan. 31, 1936, Taketani
elucidates the stereo-structural logic of quantum mechanics and analyses the problem of observation.

First, Taketani points out that “mathematics and physics learned such various facts from nature that could hardly be veiled by the formal thought”. “There has arisen in consequence the so-called mathematization, which insists that “equation is every thing!” He criticizes that, though “it has been emphasized that the method of physics should be the formal logic or analytical logic”, “this is based on a superficial understanding of physics”.

“It is in a deeper dialectical comprehension of essence and phenomenon that every achievement in physics consists”. “The interconnection between essence and phenomenon leads us directly to the comprehension of that between observation and cognition”. Namely, Taketani says that “without observation cognition cannot be obtained, but observation itself is not cognition”, and “from the confusion of these two facts there arise troubles”. “Heisenberg’s uncertainty principle indicates the limitation to each observation, but never the limitation of cognition”.

Grasping the meaning of cognition in physics in this way, Taketani begins his analysis of the logical structure of quantum mechanics with the following words — “In quantum mechanics, the phenomenal forms of wave and particle contradicting with each other, that is, the two images which the understanding excludes from each other, are grasped through their unification into the essential concept of state”. The concept of superposition of states is both very important and fundamental. “When a state is formed from a number of states by superposition, Umbestimmtheit (indeterminacy) prevails among the states so superposed. In observation, there arises the reduction of states through which Bestimmtheit (determinacy) becomes prevailing for the first time”.

“To which state it transits through the reduction can be predicted in this case only in a probability sense”. Taketani stresses, however, that “this probability is of a character quite different from a classical one”. That is, in quantum mechanics “what is unbestimmt (indeterminate) in its nature becomes bestimmt (determinate) by observation”, whereas “in classical probability every thing is bestimmt, and what is unbekannt (unknown) becomes bekannt (known) by observation”, says Taketani making use of Pauli’s wording.38

“The superposition of states is fundamentally different from that of probabilities. In other words, a wave function \( \psi \) that represents a state has a meaning more profound than probability. This is essential. That is, laws in quantum

---

38) W. Pauli, “Handbuch der Physik” (1933, springer), Bd. 24.
mechanics are not statistical ones. Probability comes out in observation. It is a phenomenal form. Since only by the abstract process of $\psi \rightarrow |\psi|^2$ is the wave function bestowed a meaning in the sense of probability, the statistical laws are nothing but phenomenological descriptions of observation. In doing so, abstraction is made of the phase factor which plays a role in interference”.

“In the kinetic theory of gas, the motion of each molecule is unbekant, but is bestimmt and obeys the law of causality. Thermal phenomena in a mass of molecules are however statistical”. In this case, the position coordinates and momentum of each molecule are called hidden parameters. On the other hand, in quantum mechanics, “the theory sets forth itself as one that can be complete without such parameters”. Taketani says that “to require such parameters is nothing but to deteriorate the necessity to the contingency”.

In this connection, Taketani takes the history of the theory of relativity, and points out that “the absolute motion with reference to the aether was a kind of hidden parameter”. “Every effort was made in its search, but ended in failure. It was at this moment that the theory of relativity came out to discard the objectively meaningless parameter, representing the world as the one in which only relative motions are meaningful. This was the first time that time, space, matter and motion were grasped in close dialectical interconnections”.

Taketani attaches importance to the fact that “in quantum mechanics, when one system is composed of two systems, there appears in the whole what is more than the mere sum of the constituent parts”. This was pointed out by Weyl in 1928 in his book “Group Theory and Quantum Mechanics”. Taketani says that “this is the relation between the part and the whole, which cannot be understood in the formal logic”.

Now, Taketani analyses the process of observation as follows. “The observation is to construct a composite system out of the observer (not the subject) and the object. The composite system as the whole is causal in its strict meaning. However, the observation is not concerned with this composition only. To be content with only this is a kind of mechanism, or an abstract viewpoint, not seeing things concretely. The observation is to extract from the whole the object, that is, a part, with the setting up of a cut. In this case, the laws of identity and of sufficient reason do not hold for the part, since it is extracted from the whole, which is more than the mere sum of the parts. As a result,

39) Weyl, “Gruppentheorie und Quantenmechanik” (S. Hirzel, Leipzig, 1928), Kap. II. Weyl says, after having treated the quantum mechanics of two particles, “In concise version, valid is in the quantum mechanics the proposition ‘the whole is more than the sum of its parts’ …”.
reduction of state takes place. That is, one and the same state appears differently in the observation, and the laws of identity and of sufficient reason do not hold because the appearance can be predicted only in terms of probability. — This is the true content of the so-called denial of causality. — It is the process of self-realization in which the essence appears different in relation to the whole.

In this way the problem, the solution of which Schrödinger came near but did not actually arrive at, in spite of his struggle with “the maximality of the catalog of expectation” brought forth by himself, it solved by Taketani independently of and nearly at the same time as Schrödinger.

Taketani raises further the question “Should we despair of getting any image of the existence on account of the limitation in observation, or other reasons?”, and answers “No, certainly not!”. “A single observation is contingent, but observations in total can reflect it as completely as possible (by making a large number of observations). — Completely does not mean correctly, but without missing. Correctness involves historical restrictions. — A sum of observations, in the sense of collecting up, constitutes statistical as well as phenomenological laws of description of phenomena; but human thoughts penetrate deeper into phenomena, thus arriving at essential as well as necessitarian laws. The Schrödinger equation has such a meaning”.

Taketani, on the relation between causality and the statistical law in quantum mechanics and in the kinetic theory of gases, points out also that “both the statistical laws of phenomena are founded on the necessitarian causal laws, though their directions are opposite to each other”. Namely, “in the kinetic theory of gases, the law of the phenomenon as a whole is statistical, being founded on the hidden parameters of each particle that obeys the causality in the strict sense”, whereas “quantum mechanics is statistical in the part, that is, in the phenomenon, being founded on the state of the system as a whole that obeys causality in the strict sense”.

Thus, Taketani makes clear that “in quantum mechanics the dialectic of unification of antagonism, that of essence and phenomenon, that of part and whole, as well as that of necessity and contingency, are closely interconnected to each other”. He says, in his paper “On Research Contributions of Prof. Niels Bohr”\(^ {40} \) received on May 27, 1937, that “this constitutes the basis for elucidation of the stereo-structure of quantum mechanics, that the probability or

\(^{40}\) M. Taketani, Sekai Bunka (World Culture, in Japanese), Jun., 1937 [the same as Ref. 18] of the present chapter; cf. also Footnote 37) above.
statistical nature is a phenomenal form while the 'state' is essential, and laws obeyed by the wave function are not statistical but essentialistic, while description by means of probability space or statistical description is concerned with their phenomenal aspects". The vagueness of the thought in the Copenhagen spirit, which we have seen before in the present chapter, is caused by a lack of understanding of this stereo-structure of the logic of quantum mechanics.

§4.2 Irreversible macroscopic process — the objective boundary between observer and object

As we have described in the last subsection §3, Schrödinger did not go beyond the application of his own argument on the system of two bodies to a system of measured object and measuring instrument, though he thought that "we must attempt to grasp objectively the interaction between the measured object and the measuring instrument". Taketani makes clear, at a symposium on the interpretation of quantum mechanics held on Oct. 14, 1942 by Gakujutsu Kenkyu Kaigi (Council for Scientific Research), that the boundary between the measured object and the measuring instrument lies between the microscopic object and the macroscopic measuring instrument in which some irreversible change takes place.\(^{41}\)

Taketani sees that "the problem of observation has been exaggerated and misunderstood to become a hot topics, on account of its remarkable results and certain tendencies of contemporary philosophy". "The tendency of philosophy in recent years has been, having abandoned epistemology in the earlier years, to aim at realism, and the problem of practice has become its main theme. Thus, the so-called action of the observer in the problem of observation in quantum mechanics has been a favorite matter for discussion of philosophers. Leaving out the structure of quantum mechanics and being encouraged by the early subjectivistic interpretation of the uncertainty principle, philosophers have been talking about the action of the subject. Such movement of philosophers has tempted, so to speak, some physicists to accept the action of the subject without trying to get a deeper understanding of the structure of quantum mechanics".

\(^{41}\)A short summary of Taketani's analysis was given in his paper "Modern Physics and Epistemology", Sizen Kagaku, Jul. 1946 [same as Ref. 9] of the present chapter. His note for the symposium was once lost, but was found later and recorded in his paper "Observation Problem of Quantum Mechanics", Kagaku (Science, in Japanese), Mar. 1964; cf. Ref. 37) of the present chapter.
Now, taking v. Neumann’s theory of observation itself, Taketani points out that v. Neumann’s theory, which we mentioned in §2 of the present section, “was aimed principally at the elucidation of the logical structure of quantum mechanics, not at the treatment of concrete and actual quantum-mechanical observations”. “This point in question was indicated by Schrödinger. Namely, he presented the famous case of a cat”. About “the paradox presented by Schrödinger”, Taketani emphasizes that “the fundamental point is that the states of life and death of the cat are not quantum-mechanical eigenstates or eigenvalues”. “Positions of the pointer of the recorder of a measuring instrument are not quantum-mechanical eigenstates or eigenvalues, too”.

“In the process that an α-particle comes into the measuring instrument to trigger it and the signal is amplified to be recorded, the state of the α-particle is meanwhile transmitted in a way different from that given by v. Neumann’s theory, or not in the form eigenvalue → eigenvalue, giving rise to a mere image of the eigenvalue of the α-particle superposed from many quantum-mechanical eigenvalues. Such are concrete and actual observations. Namely, the boundary between the measured object and the measuring instrument is not arbitrarily but is objectively determined”.

“Within the scope of v. Neumann-Pauli’s way of thinking, the separation of a beam into two beams in an experiment of the Stern-Gerlach type, for example, is considered to be an observation. In this case, the two separated beams can be superposed again to form a beam like the original one. However, it is the characteristic of an actual observation that an irreversible process is allowed to go forward, as in the case of the detection of a α-particle by a counter”. Taketani’s reference to Pauli given here is based on the following fact.

Pauli took the Stern-Gerlach experiment as a model for his “general discussion of the concept of measurement” made in his book cited above,\(^{42}\) and generalized the form of the result so obtained. In this discussion, Pauli first calculates the internal energy \(E_n(Q)\) of an atom in an external magnetic field for every value of the center of mass coordinate \(Q\) of the atom, by making use of the Schrödinger equation on the assumption that the magnetic field changes slowly in a domain comparable with the size of the atom. He supposes that the center of mass motion obeys the Schrödinger equation with the potential energy put equal to \(E_n(Q)\).\(^{43}\) It results in that the beam splits into beams

\(^{42}\)W. Pauli, Ref. 38) of the present chapter.

\(^{43}\)This corresponds to a method of calculation which is nowadays called adiabatic approximation.
one for every value of the $z$ component of the angular momentum of the atom, and the internal energy of the atom takes a specific value in every separated beam. Therefore, the direction of the center of mass motion of the atom in a separated beam makes one-to-one correspondence to the internal energy of the atom.

Then, Pauli supposes that "we can consider the center of mass a special 'measuring apparatus' and the energy of internal state $E_n$ the quantity to be measured". He supposes further that "instead of the center of mass of the atom any other apparatus could be employed provided it is realized that this apparatus certainly reacts differently to the various states $E_n$". In this case, "$Q$ describes the position of some pointer". Considering in this way an abstract measuring apparatus, Pauli states that the question "at what place one draws the cutting between the measured system described by wave functions and the measuring apparatus" is "thoroughly discussed in Chap. VI" of v. Neumann's book.

Now, Taketani continues to say, in his note for the symposium, that "in an actual observation, the reduction of state does not take place when the reading of the measuring apparatus is observed". "Also, whether the cat is alive or dead is not changed when one observes Schrödinger's cat by opening the lid of the box, though the positions of some of the elections constituting the furs of the cat may undergo reduction. That is, v. Neumann's sequence $I \rightarrow II \rightarrow III$ tends with extreme complexity towards some electrons in the furs of the cat, but not toward the state of the life of the cat".

"In the argument given above the irreversibility is essential, whereas in v. Neumann-Pauli's consideration discussions are made within the realm of reversibility. Any actual observation consists in letting an irreversible process complete itself, and an instrument of this kind is a measuring instrument".

In this way, Taketani made it clear for the first time, in a state of isolation from European and American physics and in the height of the World War II, that the quantum-mechanical measuring instrument is one in which an irreversible change is triggered by the quantum-mechanical object to be observed, and that the boundary between the object and the instrument lies between the microscopic object and the macroscopic physical system to perform the irreversible process.

Taketani says, in his 1964 paper "Observation Problem of Quantum Mechanics"\(^{44}\) that "it is because life gives us the most dramatic impression of

\(^{44}\)Cf. Footnote 41) given above.
irreversible processes that Schrödinger’s paradox of a cat causes us an essentially deep impression”, and that he “got the idea of irreversible process essentially from this case”.

3. Around the Logic of Quantum Mechanics

§1. Einstein-Podolsky-Rosen’s doubt

As we have described at the end of §4.1 of the last section, the vagueness of the thought in the so-called Copenhagen spirit is caused by a lack of understanding of the stereo-structure of the logic of quantum mechanics. Heisenberg stressed that “one must state definite experiments” to measure the position of the electron, in order to be clear about “what should be understood” by the phrase “the position of the electron”, and that the perception of the present is in principle restricted because of the uncertainty relation, so that causality is not valid. As we have seen also in §4.1 of the last section, however, one should not despair of getting any image of the existence on account of the limitation in observation, and should note that without observation cognition cannot be obtained, but observation itself is not cognition.

Bohr considered, on the basis of Heisenberg’s thought, that the space-time description and the demand for causality take on, in quantum mechanics, “complementary but to each other excluding features of description that symbolize the idealization of the possibility of observation or definition”. In spite of the fact that the uncertainty relation “indicates the limitation in each observation but never the limitation of cognition”, as is shown in Taketani’s paper cited above, the Copenhagen spirit attaches importance to the uncertainty relation and puts the phenomenological description of observation at the front.

A doubt was already raised, about eight or nine months respectively before the appearance of Schrödinger’s or Taketani’s paper cited in the last section, by Einstein, Podolsky and Rosen in their paper “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” received on Mar. 15, 1935. Einstein-Podolsky-Rosen (abbreviated EPR hereafter) put forth, for a requirement which “for a complete theory seems to be a necessary one”, the condition that “every element of the physical reality must have a counter-part in the physical theory”. They call this “the condition of completeness”.

EPR say that though "the elements of the physical reality cannot be determined by a priori philosophical considerations", they will "be satisfied with the following criterion". That is, "if, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity".

If, for example, a particle is in an eigenstate of its momentum with eigenvalue $p_0$, "the momentum has certainly the value $p_0". EPR consider that in this case "it thus has meaning to say that the momentum of the particle is real". However, in this case "a definite value of the coordinate is not predictable, but may be obtained only by a direct measurement". "Such a measurement disturbs the particle and thus alters its state". EPR say that "the usual conclusion from this in quantum mechanics is that when the momentum of a particle is known, its coordinate has no physical reality". They add that a similar statement holds generally in the case of a pair of physical quantities corresponding to non-commutable operators.

The "criterion" for "element of physical reality", which EPR suppose to be sufficient for the purpose of their paper, thus makes a clear formulation in their own way of the interpretation of quantum mechanics according to the Copenhagen spirit. It is EPR's aim to show that this criterion leads to some contradiction and the description of reality by quantum mechanics is thus not complete.

Now, EPR state that, from the point given above of a pair of physical quantities corresponding to noncommutable operators, "it follows that either

1. the quantum-mechanical description of reality given by the wave functions is not complete or
2. when the operators corresponding to two physical quantities do not commute, the two quantities cannot have simultaneous reality". "For if both of them had simultaneous reality — and thus definite values — these values would enter into the complete description according to the condition of completeness. If then the wave function provided such a complete description of reality, it would contain these values: these would then be predictable. This not being the case, we are left with the alternatives stated".

In this EPR's proof use is made of "the condition of completeness", that is, the condition that every element of the physical reality must have a counter
part in the physical theory. Though this condition itself is not false, it is a phenomenological and plane-projected view, not a stereo-structural view. As is explained in Taketani’s paper we quoted in §4.1 of the last section, “it is in a deeper dialectical comprehension of essence and phenomenon that every achievement of physics consists”, and “before one formulates equations, one must know the substantial structure in the sense of what there are and under what interaction they are, namely the model”.

“Cognition in physics is not mere phenomenology”. Certainly, “cognition is copying”. “This copying is however not a dead static reflection like that of a mirror to form the image”, and “seen historically, cognition is the process of conforming the copy to the object”. It is such a stereo-structural view of cognition that forms the core of the three-stage theory developed by Taketani.

EPR consider then the case of two systems I and II which are permitted to interact for a time interval, assuming as usual that “the wave function does contain a complete description of the physical reality of the system”. Let $\Psi(x_1, x_2)$ be the wave function of the combined system I+II left after the interaction, where $x_1$ and $x_2$ denote representatively the coordinates of the systems I and II, respectively. Let $a_1, a_2, \ldots$ be the eigenvalues of some quantity $A$ of system I, and $u_1(x_1), u_2(x_1), \ldots$ the corresponding eigenfunctions. Expanding $\Psi$ as a function of $x_1$ in terms of $u_1, u_2, \ldots$, they put

$$\Psi(x_1, x_2) = \sum_{n=1}^{\infty} \psi_n(x_2) u_n(x_1),$$

where $\psi_n(x_2)$ are regarded as the coefficients of expansion. If $A$ is measured and known to have the value $a_k$, system I is left in the state given by the wave function $u_k(x_1)$. System II is consequently left in the state given by the wave function $\psi_k(x_2)$. Choosing, instead of $A$, another quantity $B$ of system I whose eigenvalues and eigenfunctions are denoted by $b_1, b_2, \ldots$ and $\nu_1(x_1), \nu_2(x_2), \ldots$, they take, instead of Eq. (39), the expansion

$$\Psi(x_1, x_2) = \sum_{s=1}^{\infty} \varphi_n(x_2) \nu_s(x_1).$$

If $B$ is measured and known to have value $b_r$, system I is left in the state given by $\nu_r(x_1)$, and system II in the state given by $\varphi_r(x_2)$.

\footnote{$\psi_k(x_2)$ and $\varphi_r(x_2)$ are expressed as $\int u_k^*(x_1) \Psi(x_1, x_2) dx_1$ and $\int \nu_r^*(x_1) \Psi(x_1, x_2) dx_1$, respectively.}
System II is thus left in different states, as a consequence of the two different measurements performed upon system I. EPR think, on the other hand, that “since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system”. They emphasize that “this is, of course, merely a statement of what is meant by the absence of an interaction between the two systems”.

Now, EPR suppose that “it may happen that the two wave functions $\psi_n$ and $\varphi_s$ are eigenfunctions of two noncommuting operators corresponding to some physical quantities $P$ and $Q$, respectively”. They take for example the case of two one-dimensionally moving particles, and assume that $\Psi$ is given by

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \exp[ip(x_1 - x_2 + x_0)/\hbar]dp,$$

(41)

where $x_0$ is some constant. If the momentum of particle I is chosen as $A$, and is known to have the value $p$ by a measurement, particle II is left in the eigenstate corresponding to the eigenvalue $-p$ of its momentum $P$. On the other hand, if $B$ is the coordinate of particle I and is measured to have the value $x$ [the eigenfunction being $\nu_x(x_1) = \delta(x_1 - x)$], particle I is left in the eigenstate corresponding to the eigenvalue $x_1 + x_0$ of its coordinate $Q$. The operators corresponding to $P$ and $Q$ of particle II are of course noncommutable.

Then, assuming that statements of the same kind hold also in the general case considered above, EPR conclude that “by measuring either $A$ or $B$ we are in a position to predict with certainty, and without in any way disturbing the second system, either the value of the quantity $P$ (that is $p_n$) or the value of the quantity $Q$ (that is $q_s$)”. However, “both the wave functions $\psi_n$ and $\varphi_s$ belong to the same reality”. The proposition [2] given above is thus negated. They arrive in this way at the negation of [2] on the assumption that the wave function does give a complete description of the whole system, that is, on the negation of [1] which is incompatible with [2]. In consequence, EPR emphasize that “we are thus forced to conclude that the quantum-mechanical description of physical reality given by wave functions is not complete”.

EPR’s assertion is based on their thought that, after the systems I and II have ceased to interact with each other, any measurement performed upon system I could not disturb system II. However, as is clarified in Taketani’s
paper we quoted in §4.1 of the last section, in the composite system I+II "there appears in the whole what is more than the mere sum of the constituent parts", the systems I and II. The superposition expansion given in Eqs. (39) and (40) have such a meaning. Even when the two systems have ceased to interact with each other, they are not independent of each other in the sense of formal logic. EPR's discussion described above arose from their doubt of the vague interpretation of quantum mechanics according to the Copenhagen spirit. But, their discussion was lacking in the understanding of the dialectic stereo-structure of the logic of quantum mechanics. (On this point the reader is referred further to §5 of the present section, p. 299.)

§2. Bohr's 'reply' to EPR

Bohr replied to EPR in his paper received on Jul. 13, 1935,\(^\text{48}\) about three months and a half later than EPR's paper, with the same title as that of EPR's paper, and attempted to "explain a viewpoint termed 'complementarity' from which quantum-mechanical description of physical phenomena would seem to fulfill, within its scope, all rational demands of completeness".

In this paper, Bohr discusses the uncertainty relation in measurements again with the use of the wording of complementarity, and stresses that "the dependence on the reference system, in the relativity theory, of all readings of scales and clocks may even be compared with the essentially uncontrollable exchange of momentum or energy between the objects of measurements and all instruments defining the space-time of reference, which in quantum theory confronts us with the situation characterized by the notion of complementarity".

That is, Bohr asserts complementarity to be a principle like the principle of relativity, saying that "in fact this new feature of natural philosophy means a radical revision of our attitude as regards physical reality, which may be paralleled with the fundamental modification of all ideas regarding the absolute character of physical phenomena, brought about by the general theory of relativity". Bohr thinks, in other words, that EPR's assertion comes from a lack of understanding of complementarity which should be compared with the principle of relativity, on the side of Einstein, the discoverer of the very principle — in spite of the fact that EPR's doubt was presented because of the vagueness of the interpretation of quantum mechanics in accordance with the Copenhagen spirit.

\(^{48}\)N. Bohr, Phys. Rev. 48 (1935), 696.
According to Bohr, "we have in each experimental arrangement suited for the study of proper quantum phenomena not merely to do with an ignorance of the value of certain physical quantities, but with the impossibility of defining these quantities in an unambiguous way". He says that "the last remarks apply equally well to the special problem treated by Einstein, Podolsky and Rosen", and considers that "the particular quantum-mechanical state of two free particles, for which they give an explicit mathematical expression, may be reproduced, at least in principle, by a simple experimental arrangement".

This "experimental arrangement" means that "comprising a rigid diaphragm with two parallel slits, which are very narrow compared with their separation, and through each of which one particle with a given initial momentum passes independently of the other". The particles show diffraction phenomena when they pass through the slits. Bohr says that "if the momentum of this diaphragm is measured accurately, before as well as after the passing of the particles, we shall in fact know the sum of the components perpendicular to the slits of the momenta of the two escaping particles, as well as the difference of their initial positional coordinates in the same direction". He concludes accordingly that a subsequent single measurement of one of the particles "will automatically determine the position or momentum, respectively, of the other particle".

Though the wave function of "one particle with a given initial momentum" is represented by a plane wave, Bohr makes no mention of the means of letting the particle pass through one of the two slits. Anyway, Bohr emphasizes by partly using italics that "we are, in the 'freedom of choice' offered by the last arrangement, just concerned with a discrimination between different experimental procedures which allow of the unambiguous use of complementary classical concepts".

According to Bohr, because the position of one of the particles is determined by means of a correlation between its behavior and some instrument rigidly fixed to the support, which defines the space frame of reference, one loses the basis for an unambiguous application of the idea of momentum in predictions regarding the behavior of the second particle, when one allows an uncontrollable momentum to pass from the first particle into the support, by letting the first particle pass through one of the two slits. And, conversely, if one chooses to measure the momentum of one of the particles, one has no basis for predictions regarding the location of the other particle. For this reason,
Bohr supposes the existence of the freedom of “a discrimination between different experimental procedures” described above.

About the case treated by EPR, Bohr says that, though “of course there is no question of a mechanical disturbance of the system under investigation during the last critical stage of the measuring procedure”, “even at this stage there is essentially the question of an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system”. He thinks that “since these conditions constitute an inherent element of the description of any phenomenon to which the term ‘physical reality’ can be properly attached”, EPR’s discussion “contains an ambiguity as regards the meaning of the expression ‘without disturbing in any way a system’”. Thus, Bohr asserts that “the argumentation of the mentioned authors does not justify their conclusion that quantum-mechanical description is essentially incomplete”.

In this way, Bohr classifies the freedom of choice in measurement as to applying which of the two EPR’s expansions, (39) and (40), of the wave function of the composite system into the category of “inherent element of the description of any phenomenon”. According to Bohr, there is no strict separation between the object (the observed object) and the subject (the observer), and it is “a question of appropriateness” which place between them one introduces the boundary into, as we have seen before in Sec. 1, §5 of the present chapter. Bohr now supposes in addition the choice of the object to be measured itself is an element of the constitution of our theory. Bohr’s argument given in his paper under consideration thus makes the interpretation of quantum mechanics in accordance with the Copenhagen spirit more vague, so that it is incapable of replying to EPR’s paper.

By the way, about “the place where the discrimination is made between object and agencies”, Bohr states in this paper that “we have only a free choice of this place within a region where the quantum-mechanical description of the process concerned is effectively equivalent with the classical description”. No condition is shown explicitly of “effective equivalence with the classical description”. However, Bohr says that in the case that the diaphragm is not rigidly fixed to the support, it “must be treated as an object of investigation, in the sense that the quantum-mechanical uncertain relations must be taken explicitly into account”, but “if it is sufficiently massive in comparison with the particle, ... it, after the first determination of its momentum, will remain at rest”. This implies that the largeness of mass is the condition concerned.
Also, Bohr states that “the singular position of measuring instruments in the account of quantum phenomena, appears closely analogous to the well-known necessity in relativity theory of upholding a sharp distinction between space and time, although the very essence of this theory is the establishment of new physical laws, in the comprehension of which we must renounce the customary separation of space and time ideas”. In other words, Bohr seeks an analogy of the fact that the law of motion is expressed in the same form in any system of space-time coordinates, in spite of the intermixing of space- and time-coordinates on account of the Lorentz transformation, to the view that the boundary between the object to be measured and the measuring instrument is arbitrary. That this view is false has been shown by Taketani, as we described in §4 of the last section.

§3. **Furry’s ‘support’ to Bohr**

Furry supported, in his paper “Note on the Quantum-Mechanical Theory of Measurement” received on Nov. 12, 1935, Bohr’s assertion shown above by saying “Bohr has remarked that one must be careful not to suppose that a system is an independent seat of ‘real’ attributes simply because it has ceased to interact dynamically with other systems”. On the other hand, seeing that EPR’s paper “has shown the sort of situations in which this characteristic of quantum mechanics may become especially prominent”, Furry considered that “this indicates an extension of the usual discussions of the theory of measurement”.

In this paper Furry considers, in relation to the criterion proposed by EPR, “the assumption that a system once freed from dynamical interference can be regarded as possessing independently real properties”. He gives “a definite form to this assumption”, stating the following “Assumption and method A”.

“Assumption and method A: We assume that during the interaction of the two systems each system made a transition to a definite state, in which it now is. These transitions are not causally determined, and there is no way of finding out which transitions occurred, except by making a suitable measurement. In the absence of measurements we know only that the probabilities of the different transitions are respectively \( u_k \), and that if system I is in the state \( \varphi_{\lambda_k} \) system II is in the state \( \xi_{\rho_k} \). \( (k = 1, 2, 3, \ldots) \).”

---

49) W. H. Furry, Phys. Rev. 49 (1936), 393.

50) Here, the state represented by the wave function \( \varphi \) is simply called the state \( \varphi \).
With “Method B”, Furry makes “quantum-mechanical calculations”. But, in Furry’s “quantum-mechanical calculations” use is made of Eq. (39) in §1 of the present section, which was used also by EPR, and Eq. (34) in §2 of the last section introduced by v. Neumann. They are, written in Furry’s notations,

\[ \Psi(x_1, x_2) = \sum_\mu \psi_\mu(x_1) \zeta_\mu(x_2), \]  

\[ \Psi(x_1, x_2) = \sum_\mu \sqrt{w_k} \varphi_\lambda_k(x_1) \xi_\rho_k(x_2), \]

respectively.\(^{51}\)

In Eq. (42), \( \psi_\mu(x_1) \) are eigenfunctions of a quantity \( M \) of system I corresponding to eigenvalues \( \mu \). This equation means that if a value \( \mu \) is obtained in a measurement of \( M \), system II is left in the state \( \zeta_\mu(x_2) \). And then, if a quantity \( S \) of system II is measured, the probability of getting value \( \sigma \) of \( S \) is given by the absolute square of \( (\eta_\sigma, \zeta_\mu) \), the inner product of \( \zeta_\mu(x_2) \) with \( \eta_\sigma(x) \), \( \sigma \) and \( \eta_\sigma(x_2) \) being eigenvalues and eigenfunctions of \( S \) respectively. These statements applied by Furry are indeed “quantum-mechanical” as he says.

Equation (43) is, on the other hand, very particular to v. Neumann’s theory of measurement. The \( \varphi_\lambda_k(x_1) \) are the eigenfunctions of a quantity \( L \) of system I corresponding to eigenvalues \( \lambda_k \), and \( \xi_\rho_k(x_2) \) the eigenfunctions of a quantity \( R \) of system II corresponding to eigenvalues \( \rho_k \), which are supposed to make one-to-one correspondence to \( \lambda_k \). As is noted also by Furry, the quantity \( L \) of system I “serves as a ‘pointer reading’”, so that a measurement of \( L \) suffices to give the value of the quantity \( R \) of system II. (System I and II are named in Furry’s paper inversely to those in v. Neumann’s book.)

Furry says, referring to v. Neumann’s book cited before, that the expansion (43) “always exists and is in general unique”. However, as is clear from what we have described in §2 of the last section, Eq. (43) is derived from the premise, implicitly introduced by v. Neumann, that the expectation value of \( R \) is the same both in the aggregate (the mixture) of the composite system I+II and in that of system II. (The same statement applies to I+II and I.) This premise is a critical condition for v. Neumann’s theory of measurement, but is not in general valid. No proof of this condition was given by v. Neumann. Furry’s “Method B” cannot therefore be seen literally as a way of “quantum-mechanical calculations”. It should rather be said to be that of calculations

\(^{51}\) \( \zeta_\mu(x_2) = \int \psi_\mu^*(x_1) \Psi(x_1, x_2) dx_1. \) Cf. Footnote 46) in the present section.
According to v. Neumann's theory of measurement.

Now, Furry calculates probabilities in the following four cases with Methods A and B.

(a) The probability of obtaining the result $\sigma$, if no measurement is made on system I and $S$ is measured on system II. Both Methods give the same result, $\sum_k w_k |(\xi_{\rho_k}, \eta_\sigma)|^2$.

(b) The probability of finding the value $\sigma$ for $S$ in system II, if $L$ has been measured on system I and the value $\lambda_k$ obtained. Both Methods give the answer $|(\xi_{\rho_k}, \eta_\sigma)|^2$. In particular, when $S = R$ and $\sigma = \rho_k$, the probability becomes $|(\xi_{\rho_k}, \xi_{\rho_k'})|^2 = \delta_{k,k'}$, so that "a definite result is predicted". Furry says that "the possibility of such definite predictions was taken by EPR as a 'criterion of the physical reality' of the observable $R$; it is, par excellence, the bit of evidence which might incline one to believe Assumption A to be true". However, since use was not made of "Assumption and method A" by EPR, this statement of Furry cannot be appropriate.

(c) The Probability of finding the value $\rho_k$ for $R$ in system II, if $M$ has been measured on system I and the value $\mu$ obtained. Both Methods give a result proportional to $w_k |(\varphi_{\lambda_k}, \psi_\mu)|^2$.

(d) The probability of obtaining the value $\sigma$ for $S$ in system II, if $M$ has been measured on system I and the value $\mu$ is obtained. Method A gives a result proportional to $\sum_k w_k |(\varphi_{\lambda_k}, \psi_\mu)|^2 |(\xi_{\rho_k}, \eta_\sigma)|^2$ and Method B gives a result proportional to $|\sum_k \sqrt{w_k} (\varphi_{\lambda_k}, \psi_\mu)(\xi_{\rho_k}, \eta_\sigma)|^2$. Furry says that "the difference between these results comes from the well-known 'interference' between probability amplitudes".

Accordingly, Furry concludes that, if one adopts "the assumption that a system once freed from dynamical interference can be regarded as possessing independently real properties", and gives "a definite form to this assumption" to adopt "Assumption and method A", one gets "to certain types of questions the same answers as in quantum mechanics", but finds that "there exists a general class of cases in which contradictions occur". He says that "this conclusion means that a system and the means used to observe it are to be regarded as related in a more subtle and intimate way than was assumed in classical theory".

"It does not mean that quantum mechanics is not to be regarded as a satisfactory way of correlating and describing experience; it does illustrate the difficulty, often remarked upon by Bohr, which is inherent in the problem of the distinction between subject and object". Saying so, Furry supports
Bohr. However, since EPR did not use Furry's "Assumption and method A" and v. Neumann's Eq. (43), but did the expansion (42) only according to quantum mechanics, Furry's conclusion could not be a support to Bohr, nor an objection to EPR. From such a phenomenological viewpoint of "correlating and describing experience", as is shown by Furry's words, the problem in EPR's argument we pointed out in §1 of the present section could not be seen.

§4. Irreversible processes

§4.1 Green's model

Green asserted, in his paper "Observation in Quantum Mechanics"\textsuperscript{52} received on Jun. 9, 1958, that "the observation of an individual microscopic system is possible only by means of its interaction with a macroscopic system in a metastable state". It is meant here by "in a metastable state" that "even a small perturbation of the right kind is sufficient to precipitate a change of state readily observable on a macroscopic scale". According to him, "clouds chambers and bubble chambers contain matter which is literally in a metastable state; photographic emulsions and electron multipliers are easily seen to be metastable in a closely related sense".

Green was not the first who pointed out after World War II that an irreversible process is essential for any measuring instrument. Typically, Jordan said in his paper "On the Process of Measurement in Quantum Mechanics"\textsuperscript{53} published in Oct. 1949, that "the microphysical object of observation makes a track of macrophysical dimensions", and that "usually this is made possible by an avalanche process". He stated also that "thermodynamics is involved in quantum mechanical observation", and "this is in harmony with a fact showing irreversibility to be connected with observation". The special characteristic of Green's paper lies in the presentation of a concrete model which is mathematically treatable in a certain degree of measuring instruments making irreversible process.

Jordan wrote that "Margenau's article strongly influenced" his discussion. It refers to Margenau's paper "Critical Points in Modern Physical Theory"\textsuperscript{54} published in July 1937. In his paper, Margenau took as an example the measurement of the polarization of a photon by letting it pass through two Nicoll

\textsuperscript{52}H. S. Green, Nuovo Cimento \textbf{9} (1958), 880

\textsuperscript{53}P. Jordan, Phil. of Sci. \textbf{16} (1949), 269.

\textsuperscript{54}H. Margenau, Phil. of Sci. \textbf{4} (1937), 337.
prisms, and discussed as follows. “For this purpose I send it through a Nicoll prism, an analyzer. If it passes through I have made the measurement, for I then know its plane of polarization to be the plane in which the Nicoll transmits. But if I station another analyzer immediately behind the first, with its axis parallel to that of the first, I am absolutely certain that it will go through the second Nicoll also. Hence the first measurement has produced an eigenstate with respect to the photon’s state of polarization”.

“The error lies in forgetting that I have not made a measurement until I have ascertained, by catching the photon on a plate, that it has gone through the Nicoll”. According to Margenau, the separation of photons of a definite state of polarization by a Nicoll prism is “not the measurement”, but “the preparation of state”. To Margenau’s thought Jordan added, saying “according to my own idea”, “it is necessary for each observation to make — by a real physical process — the states of polarization incoherent with each other”.

This shows that the synthesis of Margenau’s idea (1937) and Jordan’s idea (1949) is equivalent to Taketani’s thought (1942), as described in Sec. 2, §4.2 of the present chapter, that the separation of a beam into two components in an experiment of Stern-Gerlach type is not a measurement, and “it is the characteristic of an actual measuring instrument that an irreversible process is allowed to go forward”.

However, there is a definite difference between the ideas of Margenau and Taketani. That is, Margenau negated the acausal change of state in a measurement. As we have seen in Sec. 2, §4.1 of the present chapter, Taketani made clear, in his 1936 paper, one year before Margenau, the logic that in a measurement we must first form a composite system out of the measured object and the measuring system, and then extract from the whole the object, a part, by carrying out a cutting. Since, in quantum mechanics, by this composition there would arise in the whole what is more than the mere sum of the parts, the reduction of states takes place in the measurement.

Margenau did not grasp this logic. Jordan wrote, “In this respect, my view differs from Margenau’s that the notion of ‘quantum jump’ not contained in the Schrödinger equation is indeed necessary and unavoidable”. But, Jordan did not come to find such logic as was made clear by Taketani. It was for the purpose of getting “the solution of the apparent dilemma” in theories of measurement of the Neumann type that Margenau negated acausal changes in measurements.
Namely, according to Margenau, though a measurement is “the result of an interaction between the system and some physical apparatus”, the “state function” will “change abruptly and indeterminately” if the process of interaction is regarded as a measurement, but it will “develop continuously” in accordance with the Schrödinger equation if the process is not regarded as a measurement, or if “one, while day-dreaming, fails to take conscious cognizance of the registration”. Margenau considered that this paradox could be solved by thinking that “there are no acausal transitions of states”. This thought also lacks comprehension of the stereo-structural logic of quantum mechanics made clear by Taketani.

Now, let us return to Green’s model in which the detector is assumed to be “a system of harmonic oscillators in a thermodynamically metastable state”. They divide into two groups of “x-mode” and “y-mode”, the former corresponding to a temperature $T$ and the latter to zero absolute temperature. The observed object is taken to be a particle of spin 1/2. On account of its perturbation, the $x$- and $y$-modes are coupled and there takes place “an irreversible transition in the direction of stable equilibrium”, which “can be detected by a temperature change”. The beam of particles is “spatially separated” so that those of up-spin enter into a detector $D_+$ and those of down-spin into a detector $D_-$. The $x$- and $y$-modes are assumed to have potential energies of the forms $\sum_{ij} U_{ij} x_i x_j$ and $\sum_{ij} U_{ij} y_i y_j$, respectively, where the $x_i$'s and $y_i$'s are the displacements of the oscillators of the $x$- and $y$-modes respectively ($i = 1, 2, 3, \ldots$). The interaction energy of the particle and the oscillators is supposed to be given by $-i \sum_i w_i (x_i \partial/\partial y_i - y_i \partial/\partial x_i)$ if it enters $D_+$, and by its complex conjugate if it enters $D_-$, where $w_i$’s are constants. Green calculates the “statistical matrix”, that is the matrix of the thermodynamical statistical operator of the oscillators, by making use of the Bloch equation.

Both the statistical matrices $\rho_+$ and $\rho_-$, respectively, of $D_+$ and $D_-$ take the form $\rho_T(x, x')\rho_0(y, y')$, where $x$, etc. denote $x_1, x_2, \ldots$, etc., and $\rho_T$ is the statistical matrix at the temperature $T$.

---

55) Denoting the Hamiltonian of the oscillators by $H$, its eigenvalues and eigenfunctions by $E_i$ and $\Psi_i$, respectively, and the free energy by $F$, and putting $\beta = 1/kT$, one has the thermodynamical statistical operator $\rho = \sum_i \exp(\beta(F - E_i))P\Psi_i$, $P$ being the statistical operator [cf. Footnote 26) in Sec. 2, §1 of the present chapter]. Putting $\rho = \exp(\beta F)\sigma$, one sees that $\sigma$ obeys the Bloch equation $\partial\sigma/\partial\beta = -H\sigma$. Green calculates $\sigma$ at $\beta \neq 0$ by considering that, for $\beta \to 0 (T \to \infty)$, $\sigma$ tends to $\sum_i P\Psi_i$. 

---
The wave function of the particle is assumed to be given at first by the superposition of the states of up- and down-spins, \( \psi = \psi_+ + \psi_- \). At the time \( t \) after the particle has begun to interact with the detector, the statistical matrix of the whole system is given by \( P = e^{-iH't}\rho_+\rho_-\psi^*e^{iH't} \), where \( H' \) is the interaction Hamiltonian. Green decomposes \( P \) into terms \( P_+, P_- \) and \( P_\pm \), which contain respectively \( \psi_+\psi_+^*, \psi_-\psi_-^* \) and \( \psi_+\psi_-^* \) as well as \( \psi_-\psi_+^* \). He calculates the quantities corresponding to their traces (the diagonal sums) taken with respect to the oscillators, and calls this calculation an "averaging operation", denoting it by \( \langle \langle \cdots \rangle \rangle \). As a result, he gets

\[
\langle \langle P_+ \rangle \rangle = \psi_+\psi_+^*, \quad \langle \langle P_- \rangle \rangle = \psi_-\psi_-^*, \quad \langle \langle P_\pm \rangle \rangle = (\psi_+\psi_-^* + \psi_-\psi_+^*)\Lambda \quad (\Lambda \to 0 \text{ as } N \to \infty),
\]

where \( N \) is the number of the oscillators.

Green concludes thus that "states in which the spin is indeterminate are rapidly suppressed". No mention is made by him of the question whether such a conclusion can be deduced from \( \langle \langle P_\pm \rangle \rangle \) obtained by the "averaging operation". Even if that is the case, Green's calculation does not prove the so-called reduction of wave function in a measurement. This is because Green makes use of the thermodynamical statistical operator.\(^{56}\)

As we have seen in Sec. 2, §1 of the present chapter, the thermodynamical statistical operator was the same one that was introduced by v. Neumann to represent the statistical character of an aggregate of a large number of particles, in which the number of particles in the state \( \Psi_i \) is proportional to the Boltzmann factor \( \exp(-E_i/kT) \). The Boltzmann factor is proportional to the probability of finding particles having the energy \( E_i \) in an aggregate at thermodynamical equilibrium corresponding to temperature \( T \). For this aggregate to arrive at the thermodynamical equilibrium, certain irreversible processes have taken place. For this reason, in the thermodynamical statistical operator one has probabilities, but not probability amplitudes themselves.

Generally speaking, from the fact that, when a physical quantity \( R \) is measured on an aggregate of systems staying in the state \( \varphi \), quantum mechanics demands that each system makes an acausal transition and is left in one of the eigenstates of \( R \), say \( \varphi_n \), with the probability \( w_n = |\langle \varphi_n, \varphi \rangle|^2 \). v. Neumann

\(^{56}\)The reader may see, remembering Eq. (15) of the present chapter with \( R \) set to be unity, that Green's \( \langle \langle P \rangle \rangle \) gives in some sense an effective statistical operator for the particle which is under influence of the oscillators.
introduced the statistical operator $U = \sum_n w_n P_{\varphi_n}$, where $P_{\varphi_n}$ is the projection operator onto the state $\varphi_n$ [cf. Eq. (24) in Sec. 1, §1 of the present chapter]. This acausal change in its state is of course irreversible. The statistical operator deals with a "mixture" of the states after the "reduction of wave packet" has occurred, but does not deal with the superposition of states. That is, in the statistical operator the "reduction of wave packet" is already taken into account as a basic concept.

Therefore, no proof of the "reduction of state" with the use of statistical operator is logically possible. If it is allowed to deduce Green's conclusion from the result obtained by him, his paper may have a significance in the presentation of an idea of a mathematically treatable, concrete model about measuring instrument to make an irreversible change. In v. Neumann's theory of measurement, as we have described in Sec. 2, §2 and Sec. 3, §3 of the present chapter, a condition for measuring instruments of his type is imposed on the statistical operator such that the expectation value of a physical quantity $R$ of a system $I$ takes the same value, both in an aggregate of $I$'s and in an aggregate of systems composed of $I$ and another system $II$ [cf. Eq. (27) of the present chapter]. Namely, an abstract model of measuring instruments of his own type, in which the eigenvalues of $I$ and those of $II$ are premised to make a one-to-one correspondence in a statistical aggregate, is represented with the use of the statistical operator.

§4.2 About the so-called "negative measurement"

Let us add here a supplementary remark about an assertion, made recently in relation to the theory of measurement in quantum mechanics, that the theory of irreversible multiplication process "has died from the fatal wound of the paradox of negative measurement". The theory of irreversible process refers to such a thought as was presented in Taketani's analysis, Jordan's view and Green's model we have described before. On the other hand, "the paradox of negative experiment" means the following. In a measurement of a particle of spin 1/2, by means of an experimental method of the Stern-Gerlach type, for example, as in Green's model, one puts a detector $D_0$ before a splitting magnet and a detector $D_+$ only for one of the split beams. In the case that $D_0$ responds but $D_+$ does not, one sees that the particle is in the other split beam. In spite of the fact that the response of $D_+$ is "no", that is, no irreversible process has taken place, one has to conclude that the "reduction of wave packet has happened", confronting a crucial contradiction.
However, for what reason can one affirm in this case that the particle is in the other split beam which does not pass through $D_+$? One can affirm this if, and only if, it is ascertained that in this experiment the incident beam consists of particles of spin 1/2 and the splitting magnet works perfectly, so that the beam is surely separated into two beams in accordance with the directions of the spin.\footnote{\textup{It is needless to say that the detectors must work perfectly.}} For this purpose, one must put a detector $D_-$ also in the other split beam, and confirm beforehand that $D_-$ does not respond when $D_0$ and $D_+$ do, and $D_+$ does not if $D_0$ and $D_-$ do.

In fact, this is the essential point of an experiment of the Stern-Gerlach type. Only with this experimental confirmation accomplished, the case that $D_0$ responds but $D_+$ does not can be deduced to be equivalent to the case that $D_0$ and $D_-$ respond but $D_+$ does not. The so-called "negative measurement" is essentially not a measurement. $D_+$ of the "negative measurement" may serve, in an experiment of a different kind, as a monitor, guaranteed to respond anticoncurrently with $D_-$ by preliminary measurements in the preparatory stage of the experiment, for informing that the other split beam is of particles of down-spin by its absence of response. But, for the whole system of instrument, a confirmation of the kind mentioned above must necessarily be carried out in some way by means of detectors based on irreversible processes. Unless this is done, the experiment cannot be considered a measuring experiment.

Let us add a few more words about the basic theory of particle detection. Consider the simple case that there is a particle detector. When the detector responds, it is known that a particle appropriate for the detector has passed through the detector by virtue of an irreversible process. Now, when the detector does not respond, it is concluded, within the bound of the efficiency of the detector, that any particle has not passed through the detector. This is not a so-called "negative measurement", nor an event to negate the irreversible process in the measurement.

This is because the conclusion mentioned above that any particle has not passed through the detector can be deduced only on the basis of the guarantee, accomplished through a sufficient inspection, that when a particle passes through the detector the irreversible process concerned takes place. Without this guarantee nothing can be said. Namely, one sees that (a) the irreversible process takes place when a particle comes into the detector, and (b) any particle has not come into the detector when the irreversible process does not take place, both the cases being connected with each other on the premise
of the irreversible process. Thus, a measuring device based on a certain irreversible process is indispensable for the detection or no detection of a particle. One should grasp these things explained above as an elementary knowledge of experiments.

As we have described several times in the present volume, the measurement of a physical quantity of a system makes the state of the system change irreversibly to an eigenstate of the quantity. This is one of the fundamental laws of quantum mechanics. About this law, Dirac discusses in his famous text book\(^{58}\) as follows:

"When we measure a real dynamical variable \(\xi\), the disturbance involved in the act of measurement causes a jump in the state of the dynamical system. From physical continuity, if we make a second measurement of the same dynamical variable \(\xi\) immediately after the first, the result of the second measurement must be the same as that of the first. Thus after the first measurement has been made, there is no indeterminacy in the result of the second. Hence, after the first measurement has been made, the system is in an eigenstate of the dynamical variable, the eigenvalue it belongs to being equal to the result of the first measurement. This conclusion must still hold if the second measurement is not actually made. In this way we see that a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured".

Any quantum mechanical measurement has such a character. In addition, the quantum mechanical measurement is carried out by means of a measuring instrument containing a macroscopic many-body system, in which some irreversible change takes place. In the history of science, one sees that microscopic and atomic substances have been discovered and their attributes have been made clear, indeed by virtue of the development of scintillation counter, cloud chamber, ionization counter, and so on. One should not forget that quantum mechanics is the theory in the essentialistic stage in the cognition of the equations of motion of those substances.

It was analyzed by Taketani in 1939, as we have described in Sec. 2, §4.1, quantum mechanics is "statistical in the part, that is, in the phenomenon, being founded on the state of the system as a whole that obeys causality in the strict sense", while in the kinetic theory of gas "the law of the phenomenon as a whole is statistical, being founded on the hidden parameters of each particle that obeys causality in the strict sense". In the unification of both "the statistical

\(^{58}\)P. A. M. Dirac, "The Principles of Quantum Mechanics" (Cambridge; 1930, 1934, 1947).
laws of phenomena of directions opposite each other" the logical ground of quantum mechanical measuring instruments lies.

§5. "Each photon interferes only with itself" — Dirac's expression

Without having grasped the stereo-structural logic of the causal and statistical natures of quantum mechanics, which we have just described above, Heisenberg attempted to interpret the occurrence of probabilities in quantum mechanics in connection with "many atoms" from the viewpoint of the uncertainty relation, as we have seen in Sec. 1, §3 of the present chapter. The probability in quantum mechanics concerns one particle or one system, as we have pointed out several times before in the present chapter. About this point, Dirac says in his textbook cited above as follows:

"Sometime before the discovery of quantum mechanics people realized that connection between light waves and photons must be of a statistical character. What they did not clearly realize, however, was that the wave function gives information about the probability of one photon being in a particular place and not the probable number of photons in that state. The importance of the distinction can be made clear in the following way. Suppose we have a beam of light consisting of a large number of photons split up into two components of equal intensity. On the assumption that the intensity of a beam is connected with the probable number of photons in it, we should have half the total number of photons going into each component".

"If the two components are now made to interfere, we should require a photon in one component to be able to interfere with one in the other. Sometimes these two photons would have to annihilate one another and other times they would have to produce four photons. This would contradict the conservation of energy. The new theory, which connects the wave function with probabilities for one photon, gets over the difficulty by making each photon go partly into each of the two components. Each photon then interferes only with itself. Interference between two different photons never occur".

Now, by virtue of the development of laser technology, it has become possible with a full application of electronics to perform experiments of the interference between two beams of light coming from two laser sources, instead of that between two separated components of a beam of light coming from a source of light as in the Young experiment. Magyar and Mandel\(^59\) observed

interference effects by superposing beams of light from two laser sources. Mandel\(^{60}\) thought that a question arose as to the relation between their result and Dirac's expression "each photon interferes only with itself. Interference between two different photons never occur".

Pfleegor and Mandel\(^{61}\) thus made an experiment "under conditions where the light intensity was so low that the mean interval between photons was great compared with their transit time through the apparatus". They say that "loosely speaking, it could be said that, with high probability, one photon was absorbed sometime before the next one was emitted by one or the other source". It is their result that "even under these conditions interference effects are observable in the superposition of two independent light beams".

Pfleegor-Mandel argue that "whereas in a conventional interferometer it is possible (if undesirable) to think of each photon going partly into each beam, this point of view is even less valid here, since the beams originate in two separate sources", and say that "it seems better to associate the interference with the detection process itself, in the sense that the localization of a photon at the detector makes it intrinsically uncertain from which of the two source it came".

In other words, if each photon is supposed to be a wave packet concentrating into an interference fringe, the direction from which it come is utterly unknown on account of the uncertainty relation. Pfleegor-Mandel conclude therefore that since "no measurement of the sources will yield this information without at the same time destroying the fringes", whereas "the beams originate in two separate sources" and "one photon is absorbed sometime before the next one is emitted by one or the other source", "the interference effect is associated with the detection of each photon", and "surprising as it might seem, the statement of Dirac appears to be as appropriate in the context of this experiment".

de Broglie and Andrade e Silva consider, in their paper "Interpretation of a Recent Experiment on Interference",\(^{62}\) this indefinite argument of Pfleegor-Mandel to be a good evidence for de Broglie's own wave theory which he has adopted for some years. He calls his wave "the wave \(v\)".\(^{63}\) "This wave \(v\) has a very low amplitude and does not carry energy, at least not in a noticeable

\(^{63}\) L. de Broglie, "Non-Linear Wave Mechanics" (Elsevier, Amsterdam, 1960).
manner. The particle is a very small zone of highly concentrated energy incorporated in the wave, in which it constitutes a sort of generally mobile singularity. He presented already in 1927 "the double solution theory", a theory similar to the theory of wave $v$, in which continuous solutions satisfying a wave equation and singular solutions representing particles were supposed.

Applying the theory of wave $v$ to the special case of photons, de Broglie "considers the wave $v$ of the photons as a very weak electromagnetic wave obeying the Maxwell equations very closely". About Pfleegor-Mandel's experiment, de Broglie-Andrade e Silva say that "there exist interference fringes in the apparatus even if there is no photon that permits one to detect them" on account of the wave $v$. According to them, "a photon coming from one laser or the other and arriving in the interference zone is guided by the superposition of the waves emitted by the two lasers, and it is for this reason that it is impossible to know in which of the two lasers it was born". "The part played by the photons is only to permit one to detect the interferences by the manner in which they are statistically distributed in the zone where these interferences exist".

As is pointed out by Nagasaki in his paper "No 'Need of Reexamination' of the Dual Nature of Photon", the core of the problem lies in the fact that any detection of photons was not made between one or the other laser source and the interference detector in Pfleegor-Mandel's experiment. In their experiment, they observed photons for the first time by the interference detector. In the Young interference experiment discussed by Dirac, no detection is made of photons between one or the other slit and the screen. Every photon is detected for the first time on the screen.

If the detection of photons is made between one or the other slit and the screen, the slit from which each photon comes is determined, but at the same time the phase of each photon as a light wave becomes utterly undetermined, so that no interference effect takes place on the screen. Since the two slits are, seen from the screen, the immediate sources of light, one has to say that each photon comes from both the slits to interfere with itself, according to Dirac's way of expression. In Pfleegor-Mandel's experiment, if the photons are detected between one or the other laser source and the interference detector, no interference effect takes place. One has to say that each photon coming from the two laser sources interferes with itself.

64) L. de Broglie, J. de Phys. et Radium 8 (1927), 225.
Though the phase of light wave from a customary source changes in very short times, the difference between the phases of the two components separated out of a light beam from a source remains constant. On the other hand, a laser source emits as an oscillator of light a beam of constant phase for a time considerably long compared with that of customary sources. The difference of the phases of the beams from two “independent” laser sources is consequently constant in a certain duration, so that by means of a full application of electronics their interference effects can be detected.

As we have described at the beginning of the present chapter, quantum mechanics is an essentialistic theory in which Heisenberg-Born-Jordan’s phenomenological matrix mechanics based on corpuscular theory and Schrödinger’s substantialistic wave mechanics are unified by Born’s statistical interpretation of the wave function as the moment. It was de Broglie’s theory of matter wave that gave birth to wave mechanics, and de Broglie’s theory itself was born out of his idea of interpreting light quanta as particles of extremely small mass, as we have seen in Sec. 2, Chap. 5, Vol. II of the present book.

de Broglie attempted, even after the establishment of quantum mechanics, to interpret quantum mechanics from a substantialistic viewpoint, proposing his theory of wave $v$ and others. For the theory of wave $v$ of such nature, Pfleegor-Mandel’s indefinite expression might seem to offer good evidence. Fully realized is the importance of grasping the dialectical stereo-structural logic of the essentialistic quantum mechanics, from the viewpoint of the three-stage theory of cognition of nature.

In the present chapter, we have analyzed the defect of the interpretation of quantum mechanics based on the so-called “Copenhagen spirit” due to Bohr and others, and made clear at the same time that EPR (Einstein-Podolsky-Rosen)’s doubt was caused by the vagueness of the “Copenhagen spirit”. We have pointed out also that, both in the “Copenhagen spirit” and in EPR’s doubt, the characteristic stereo-structural logic of quantum mechanics is not understood. EPR argued, taking for example two one-dimensionally movable particles, that “the quantum-mechanical description of reality given by the wave function is not complete”. Even nowadays, we sometimes encounter such discussions which call this example EPR’s paradox and, substituting for the original example the decay of a system into its two constituent particles of spin 1/2, stress the paradoxical feature of the example.

These discussions go as follows. Let us consider a composite particle $C$ consisting of two particles $A$ and $B$, both of which are of spin 1/2. $C$ is assumed
to be in a state of spin zero. In this case, the spins of A and B are anti-parallel
to each other, and the state of C is the superposition of the state I, in which A
is in the state of up-spin and B in that of down-spin, and the state II, in which
A is in the state of down-spin and B in that of up-spin. Now, suppose that we
measure the spin of A when A and B are at a long distance apart after C has
decayed into A and B, and detect A either in the state of up-spin or in that of
down-spin, the spin state of B is determined to be, respectively, either down
or up. Is not this in contradiction with the fact that A and B do not interact
with each other? Or, should we think that A and B communicate with each
other by a signal propagating with an infinitely high velocity?

However, the anti-parallelism of the spins of A and B in the case under
consideration derives from the naught of the spin of C, as well as from the
composition and conservation laws of angular momenta, but is not related
directly to the interaction between A and B. (In EPR's example, the composi­
tion and conservation laws of linear momenta are effective.) It is indeterminate
whether the spin state of A is up or down, because the state of C is the su­
perposition of the state I and II. It is just like the fact that in the Young
interference experiment it is indeterminate from which slit each photon comes.
By the measurement, the spin state of A is determined together with that of
B which should be anti-parallel to the former.

In other words, the state of the composite system A+B after the decay,
as well as the state of the composite system C before the decay in this case,
is determined according to the general laws in quantum mechanics. By a
particular measurement carried on the constituent system A, the state of the
composite system A+B is determined to be either the state I or II out of the
superposition of the states I and II. On account of this determination and of
the speciality in the composition of the state I or II in terms of the spin states
of A and B, the spin state of A becomes determinate together with that of B.

Also in this case, there holds true the stereo-structural logic characteristic
to quantum mechanics, that "there appears in the whole what is the mere
sum of the constituent parts", which we have described in Sec. 2, §4.1 of the
present section. Quantum mechanics is complete in its logical structure. But,
the logical structure of quantum mechanics is not such one of plane-projected
view as that of the formal logic, but is dialectical and stereo-structural. One
should not forget this in discussing quantum mechanics.
This page is intentionally left blank
Index for Volumes I, II, III

α-rays
  Detection of large-angle scattering of I203
Abraham
  Angular momentum carried away by radiation II159
  Magnetic moment of rotating spherical electron II220
Adams II81
Adiabatic hypothesis
  And correspondence principle II166
  As generalization of Wien’s displacement law II165
  Ehrenfest’s II165
Adiabatic invariant
  Indication of existence of II165
Allen II221
Andrade → Rutherford-
Andrade e Silva → de Broglie-
Angular momentum
  Bohr’s assumption of constancy of II93
  Carried away by radiation II159
  Conservation law and selection rules of II158
  g-algebraic treatment of III53
  Indication of quantum nature of (Nicholson) II77
  Of atomic body II184, 187
Properties in matrix mechanics of III29
Quantum nature in Bohr model of II92
Relativistic limit of (Sommerfeld) III126
Anomalous Zeeman effect
  And half-integer quantum number II182
  And Runge’s rule II181
  And Stern-Gerlach’s experiment II195
  Systematization by Back of II182
  Treatment by matrix mechanics of III31
Anomaly of g-factor of s-terms
  And magneto-mechanical anomaly II1783, 190, 191
  And magneto-optical anomaly II189
Anticommutation relation
  For de Broglie waves obeying F-D statistics (Jordan-Wigner) III225
Aperiodic motion
  Attempt at application of matrix mechanics to III61, 125
  Difficulty in applying quantum condition to II195
  Generalization of matrix mechanics to III174
Arrhenius I97
Arvidsson II191
Avogadro number
  Thomson's determination (by
  method of electrolysis) of II112
Azimuthal quantum number
  Sommerfeld's introduction of II123
  Two kinds of (Bohr-Coster) II206

Back
  Systematization of anomalous
    Zeeman effect II182
    → Pascen-
Balmer
  Presentation of spectral formula I63
Balmer formula
  Pauli's matrix mechanical derivation
    of III37
  Sommerfeld's derivation from the
    quantum condition of II118
  Schrödinger's wave mechanical
    derivation of III76
  Derivation using Bohr model of II88
  Ishiwara's derivation of a formula
    similar to II113
  Dirac's discussion based on \(q\)-algebra
    of III52
Barkla
  Determination of number of centers
    of X-ray scattering in atoms II183
  Study of secondary X-rays II180, 183
Barnet II191
Bates → Sucksmith-Beck II191
BKS theory → Bohr-Kramers-Slater
Black body radiation Kirchhoff's
  concept and law of I5
Bohr
  Assumption of constancy of angular
    momentum II93
  Assumption of quantum numbers
    invariance II218
  Configuration of inner-atomic
    electrons II98, 202

Determination of number of electron
  in hydrogen atom I229
Doubt on half-integral quantum
  numbers II208
Doubt on energy-momentum
  conservation II263
Frequency condition II87, 151; III20
Germ of correspondence principle
  II92
Inference of frequency of
  inner-atomic electrons II81
Interpretation of Franck-Hertz's
  experiment II102
Nonmechanical compulsion to
  lighting electrons II209, 216;
  III103
Presentation of Bohr model of atom
  II82
Presentation of correspondence
  principle II163
Reply to EPR's question III282
Support for negation of causality
  III248
Treatment of Stark and Zeeman
  effects by a pre-correspondence-
  principle II143
Use of the term "complementarity"
  III248, 249

Bohr-Coster
  Introduction of two kinds of
    azimuthal quantum numbers
    II206
Bohr-Kramers-Slater
  Idea of virtual oscillation II267, 275
  Theory of statistical conservation of
    energy-momentum II266
Bohr model
  And quantum nature of angular
    momentum II93
  And radius of hydrogen atom II84
  And the spectral formula II88
  And structures of atom and molecule
    II94
Basic assumptions of II86
Confirmation by Mosely of II99
Boltzmann → Stefan-Boltzmann
Dispersion formula II277
Introduction of concept of probability into wave mechanics III129, 152
Reversibility of transition probability III165
The Born approximation (of scattering probability) III135
Transition probability (of absorption and induced emission) III164
Born approximation
Wentzel’s application to Rutherford scattering III198
Born-Heisenberg-Jordan
Introduction of canonical transformation III27
Matrix mechanics of a system of multi-degrees of freedom III26
Matrix mechanics of a string III35
Matrix mechanics of angular momentum III29
Perturbation method in matrix mechanics III28
Relative intensity in Zeeman effect III30
Born-Jordan
Attempt to treat the electro-magnetic equations by matrix mechanics III33
Equation of motion in matrix mechanics III26
Intensified quantum condition III23
Matrix mechanics of systems of one degree of freedom III19
Matrixization of canonical equations of motion III21
Born-Jordan’s quantum condition and the commutation relation III22
Born-Kármán
Theory of specific heats of solids II24
Born-Wiener
Attempt at applying matrix mechanics to aperiodic motion III61, 189
Bose
Statistical mechanics of gas of light quanta II253
Bose-Einstein statistics II258; III152
Bothe-Geiger
Experiment negating the BKS theory II269
Boundary
Between an object and apparatus of observation III261, 274, 275
Bowen → Millikan-Brackett
Spectral series of I64
Breit
Idea of virtual orbit II275
Brillouin, L.
WKB approximation of wave equation III92
Brillouin, M.
Inner-atomic elastic waves and quantum condition II244
Burgers II166
Busch
Angular momentum carried away by radiation II159
Campbell II199
Canonical transformation
And wave mechanics (London) III167
In matrix mechanics III27
Dirac’s transformation theory based on III175
Catalán II186
Catalog of expectation
Wave function as (Schrödinger) III269
Clark-Duane
Anomalous scattering of X-rays  II239
Combination rule  
  Of amplitudes (classical)  III10  
  Of amplitudes (quantum-theoretical)  
    III10, 20  
  Of frequency (classical)  III8  
  Of frequency (quantum-theoretical)  
    III8, 20  
  Of spectral lines (Ritz)  II63
Commutation relation  
  And anticommutation relation  III225  
  And Born-Jordan’s intensified  
    quantum condition  III23  
  And Dirac’s Poisson bracket  
    equation  III47  
  And Heisenberg’s  
    quantum-theoretical quantum  
    condition  III23  
  And the uncertainty relation  III241  
  Of angular momenta  III29
Complementarity  
  Taketani’s critique of  III252  
  In Bohr’s wording  II247, 249
Complexion number  
  And combination with repetition  
    II257  
  And the derivation by Debye of the  
    Planck formula  II18  
    Planck’s  I28, II14
Compton  
  Light quanta and reflection and  
    diffraction of X-rays  II242  
  Scattering of X-rays and size of  
    electron  II235  
  Spin and the magneto-mechanical  
    anomaly  II221  
  Theory and experiment of Compton  
    effect  II236
Compton effect  
  Dirac’s Aplication of matrix  
    mechanics to  III203  
  Klein-Nishnina’s application of  
  Dirac’s electron wave equation to  
    III232  
  Tamm’s application of quantum field  
    theory to  III233  
  Gordon’s application of wave  
    mechanics to  III203  
  Interpretation according to the BKS  
    theory of  II268
Compton-Simon  
  Experiment negating the BKS  
    theory  II270
Condon  
  Commutation relation and the  
    uncertainty relation  III241  
  Conflict between field and matter  
    Indication by Taketani of  III225,  
    III235
Conjugate configuration  
  About a closed shell  II217  
  Conservation law of energy-momentum  
    Doubts by Bohr of  II263  
    Statistical  II256
Continuity equation  
  Of charge density (Schrödinger)  
    III123  
  Of four-current density (Klein,  
    Gordon)  III203
Contradiction in radiation theory  
  Indication by Einstein of  I49, II7
Copenhagen spirit  
  Cause of vagueness of  III274  
  In interpretation of quantum  
    mechanics  III252
Corpuscle  
  Name of electron (Thomson)  I115  
  Thomson’s model of atoms without  
    nucleus by means of  I106, 128,  
    140
Correspondence principle  
  And the adiabatic hypothesis  II166  
  And intensity of radiation  II169  
  And selection rule  II163  
  Germ of  II92
Presentation by Bohr of 11163
Coster
  Screening doublet terms and relativistic doublet terms 11205
Covalent bond
  Lewis and Langmuir's phenomenological introduction of 11198
  Heitler and Longon's Quantum mechanical elucidation of 11223
Crowther
  Single and multiple scatterings of $\beta$-rays 1207, 219
$\delta$-function
  Dirac's 11175
Darwin
  Attempt to introduce spin of vector type 111214
  On Dirac electron equation and fine structure constant 111232
  Theory of scattering and absorption of $\alpha$-particles 11222
Davisson-Kunsman 111261
de Broglie
  Prediction of diffraction of electrons 111251
  Supposition of small mass for light quantum 111248
  Virtual wave (phase wave) and the quantum condition 111249
de Broglie-Andrade e Silva $\nu$-wave and light 111297
de Broglie waves
  Theory of gases and fluctuation in gases in terms of 111259; 11172
Debye
  Derivation of Plank's distribution formula 11118, 261
  Hypothesis of elementary action and unharmonic oscillators 111114
  Theory of Compton effect 111239
  Theory of specific heats of solids 11123
Zeeman effect and the quantum condition 111141
de Hass 111190
Dennison
  Determination of nuclear spin of hydrogen atom 111223
Difficulty of frequency
  In classical oscillation models of spectral series 11173
  in wave equation 111178
Dirac
  Commutation relation and Poisson bracket equation 111146
  Compton effect by matrix mechanics 111204
Difficulty in matrix mechanics for many-body systems 111147
Equation of motion in $q$-numbers 111147
  Introduction of $\delta$-function 111175
  Introduction of transformation functions 111176
  matrices with continuous row and/or column 111174
Number representation of wave equation 111209
Permutation symmetry of $\psi$ for a many-body system 111148
Quantization of radiation as a system of light quanta 111210
Relativitization of matrix mechanics 111204
Relativitization of wave equation 111150, 154
Spin-orbit interaction based on Dirac's electron equation 111232
Theory of positive electron 111234
Transformation function and Born's scattering probability 111182
Transition probability of absorption and induced emission of radiation 111158
Treatment of Balmer formula by
Index for Volumes I, II, III

q-numbers III42
Treatment of equation of electron orbit by q-numbers III51
Wave equation (time-dependent) III154, 181
Dispersion formula
And Heisenberg’s quantum condition III11, 14
Born’s II277
By wave mechanics (Schrödinger) III118
Kramers’ II273
Kramers-Heisenberg’s II282; III6
Ladenberg-Reiche’s II272
Displacement law
Of spectral series II180
Wien’s I8, II8
Distribution law of black body radiation Michelson’s I13; II7
Planck’s I25; II11
Rayleigh-Jean’s I22; II12
Wien’s I16; II8
Doubleness of the electron Indication of (Pauli) II215
Drude
Determination of inneratomic electrons (by method of light dispersion) I191; II20
Dual nature of radiation
Indication by Einstein of II32
Intermediary viewpoint of II160, 241
Solution of conflict in III206
Duane
Light quantum and Bragg reflection II241
→ Clark-Dynamiden
Lenard’s model of atom by I163
Number in a hydrogen atom of I165
Their true radius and specific absorption power of cathode rays I166
Ehrenfest
Adiabatic hypothesis and the quantum condition II165
Adiabatic hypothesis III164
Hypothesis of elementary action and rotational motion II115
→ Epstein-
Ehrenfest-Kamerlingh Onnes Analysis of number of complexion II257
Einstein
Demand for a quantum-theoretical theory of radiation II156
Derivation of Planck’s distribution formula II150, 258
Fluctuation in a gas and de Broglie waves II259
Fluctuation of radiation energy II27
Fluctuation of radiation pressure II31
Hypothesis of light quantum I43
Indication of contradictions in theory of radiation II7
Indication of dual nature of radiation II32
Indication of existence of adiabatic invariants II165
Introduction of transition probability II149
Out-radiation, positive and negative in-radiation II149
Quantum-statistical mechanics of ideal gases II255
Recoil of molecule in the theory of light quantum II151
Relation between the elementary action and $e^2/c$ II42
Theory of specific heats of solids II19, 22
Einstein-Ehrenfest difficulty in interpretation of Stern-Gerlach’s experiment II193
Einstein-Podolsky-Rosen → EPR
Electron
Kelvin’s model of atoms without nucleus using I115
Kelvin’s name for electron  II116
Electromagnetic field equations
   Attempt to matrixize  III33
Electron configuration
   By Bohr  II98, 202
   By Kossel  II196
   By Stoner  II211
   Conjugate  II217
   Of rare gas atoms (Rydberg)  II198
Electron diffraction
   Elsasser’s Indication of experimental evidence for  II260
   Prediction by de Broglie of  II251
Electron specific charge
   Magnetomechanical anomaly of  II190, 213
   Zeeman’s measurement using the Zeeman effect  II92
   Thomson’s Measurement on free electrons of  II100
Elementary action
   And radius of atom of Nagaoka type  II72, 73
   And radius of atom of Thomson type  II44
   Relation to $e^2/c$ (Einstein)  II42
Elementary electric charge
   Equality of positive and negative elementary charges  II144
   Measurement by Thomson of (by gas ionization method)  II111
   Measurement by Townsend of (by ion diffusion method)  II113
   Measurement by Wilson of (by falling drop method)  II113
Elester-Geiter  I78
Elsasser
   indication of experimental evidence of wave nature of electron  II260
Energy quantum (element)
   Of resonator  I31
Entropy
   Of Planck’s resonator  I19, 25, 30;
   Of the radiation obeying Wien distribution law  I41
EPR
   Einstein-Podolsky-Rosen EPR’s question
   On interpretation of quantum mechanics  III279, 282, 299
   Problem of  III280, 282
   Bohr’s and Furry’s replies to  III282, 285
Epstein
   Equality of Sommerfeld’s and Planck’s quantum conditions  II140
   Quantum condition and Stark effect  II132
   Epstein-Ehrenfest
   Light quantum and X-ray diffraction  II244
Equation of motion
   In $q$-number (Dirac)  III47
   In matrix mechanics (Born-Jordan)  III25
Equipartition law of energy
   And Rayleigh-Jean’s distribution formula  I37; II12
Eucken  II115
Example of cat
   Schrödinger critique of v. Neumann’s theory of observation  III268
Exclusion principle
   Presentation by the name of general rule  II207, 218

Fermat principle
   And Hamilton principle  II240; II80
Fermi
   Quantization of ideal gas  II153
   Fermi-Dirac statistics  II260, III153
Field quantization
   Beginning of (Born-Heisenberg-Jordan)  III37
Index for Volumes I, II, III

Germ of (Rubinowicz, Compton) III160, 243
Of the charged de Broglie wave (by Jordan–Klein) III226
Of the de Broglie waves using anti-commutation relation (by Jordan–Wigner) III225
Of the de Broglie waves obeying the F-D statistics (by Jordan) III225
Of Dirac’s electron field plus electro-magnetic wave (by Heisenberg-Pauli) III228
Of free electromagnetic waves III227
Of radiation as a system of light quanta (by Dirac) III210
Second quantization of de Broglie waves III208, 224, 225
Fine structure
Of relativistic energy levels II127; III202, 232
Fine structure constant Introduction by Sommerfeld of II127
Fitzgerald
Mechanical model of spectral series I69
Floating magnets
Analogous property to the periodic law I110
And Thomson’s corpuscle model of atoms I109
Fluctuation
In the ideal gas II259; III72
Of radiation energy II27
Of radiation pressure II30
Of oscillation energy of string III37
Fock
Relativistic wave equation and the fine structure constant III202
Fowler II89
Franck-Hertz
Measurement of excitation energy of atoms II102
Frenkel
Derivation of spin-orbit interaction II213
Frequency
Larmor’s I96; II142
Of inneratomic electrons II181
Of de Broglie wave II248, 259; III84
Frequency condition
And relativity (Fürsternling) II231
Bohr’s II87, 151; III20
Fresnel band theory
Similarity to quantum condition III89
Furry
Support for Bohr’s reply to EPR’s question III285
Fürsternling
Frequency relation and theory of relativity II231

$g$-factor
Anomaly in s-terms of II183, 189, 191
Introduction by Landé of II182
Of multiplet terms II188
Problem of the relativistic correction of II213
Geiger
Measurement of scattering angle of α-rays by gold foil I208
→ Bothe-
→ Elester-
→ Rutherford-
Geiger-Marsden
Detection of large-angle scatterings of α-rays I207
Ghost field
Attempt to interpret wave function by III126
Giese I78
Gordon
Dirac electron equation and the fine structure III232
Relativization of wave equation III203
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goudsmit → Uhlenbeck-Green model of irreversible measuring apparatus</td>
<td>III290</td>
</tr>
<tr>
<td>Group velocity of de Broglie wave</td>
<td>II250, III71</td>
</tr>
<tr>
<td>Guiding wave in attempt to interpret wave function</td>
<td>II252, III127</td>
</tr>
<tr>
<td>Half-integer quantum number</td>
<td></td>
</tr>
<tr>
<td>And anomalous Zeeman effect</td>
<td>II182, 184, 186, 188</td>
</tr>
<tr>
<td>Bohr's doubts on</td>
<td>II208</td>
</tr>
<tr>
<td>Hamilton principal function And phase of wave</td>
<td>II184</td>
</tr>
<tr>
<td>Hamilton principle And Fermat's principle</td>
<td>II252, III80</td>
</tr>
<tr>
<td>Hamilton-Jacobi partial differential equation</td>
<td></td>
</tr>
<tr>
<td>And the wave equation at short wavelengths</td>
<td>III83</td>
</tr>
<tr>
<td>Hass Radius of an atom in Thomson model</td>
<td>II44</td>
</tr>
<tr>
<td>Heisenberg Angular momentum of atomic body</td>
<td>II184</td>
</tr>
<tr>
<td>Inter-atomic resonance and permutation symmetry of wave function</td>
<td>III143</td>
</tr>
<tr>
<td>Negation of causality from the uncertainty relation</td>
<td>III247</td>
</tr>
<tr>
<td>Quantum-theoretical law of multiplication</td>
<td>III10</td>
</tr>
<tr>
<td>Quantum-theoretical quantum condition</td>
<td>III11, 14</td>
</tr>
<tr>
<td>State probability and matrix element of canonical transformation</td>
<td>III172</td>
</tr>
<tr>
<td>Symmetry of wave function and spectral terms</td>
<td>III219</td>
</tr>
<tr>
<td>Symmetry of wave function and nuclear spin</td>
<td>III221</td>
</tr>
<tr>
<td>Thought experiment on the uncertainty relation</td>
<td>III239</td>
</tr>
<tr>
<td>→ Born-</td>
<td></td>
</tr>
<tr>
<td>→ Kramers-</td>
<td></td>
</tr>
<tr>
<td>→ Sommerfeld-</td>
<td></td>
</tr>
<tr>
<td>Heisenberg-Jordan Matrix mechanics of anomalous Zeeman effect</td>
<td>III31</td>
</tr>
<tr>
<td>Heisenberg-Pauli Quantization of Dirac’s electron field plus...</td>
<td>III228</td>
</tr>
<tr>
<td>Heisenberg's quantum condition And dispersion relation</td>
<td>III23</td>
</tr>
<tr>
<td>And the commutation relation</td>
<td>III22</td>
</tr>
<tr>
<td>Heitler-London Quantum mechanics of covalent bond</td>
<td>III223</td>
</tr>
<tr>
<td>Hertz Energy difference in X-ray L terms</td>
<td>II204</td>
</tr>
<tr>
<td>→ Franck-</td>
<td></td>
</tr>
<tr>
<td>Hidden variables Problem on the existence of</td>
<td>III247, 273</td>
</tr>
<tr>
<td>Hilbert-Neumann-Nordheim Mathematical systematization of transformation theory</td>
<td>III190</td>
</tr>
<tr>
<td>Hori</td>
<td>III222</td>
</tr>
<tr>
<td>Hund Symmetry property of molecular wave functions</td>
<td>III220</td>
</tr>
<tr>
<td>Hydrogen atom Bohr’s Application of the Bohr model to</td>
<td>II84</td>
</tr>
<tr>
<td>Pauli’s application of matrix mechanism to</td>
<td>III38</td>
</tr>
<tr>
<td>Dirac’s and Gordon’s application of the Dirac electron equation to</td>
<td>III232</td>
</tr>
<tr>
<td>Ishiwara’s application of the quantum condition to</td>
<td>II113</td>
</tr>
<tr>
<td>Sommerfeld’s application of the quantum condition to</td>
<td>II117, 125</td>
</tr>
<tr>
<td>Fock’s application of the relativistic</td>
<td></td>
</tr>
</tbody>
</table>
wave equation to III202
Schrödinger’s application of wave mechanism to III74
Dirac’s attempt to apply q-algebra to III50
Bohr’s determination of number of electron in I228
Hypothesis of action integral
Sommerfeld’s II48
Hypothesis of elementary action
And anharmonic oscillation II115
And rotational motion II115
Planck’s II47

In-radiation
And absorption and induced emission of light II149
Inneratomic elastic wave
And the quantum condition (Brillouin) II245
Intensiﬁed quantum condition
Born-Jordan’s (commutation relation) III24
Intensity of radiation
From ring of electrons (Schott) I197, 206
Of Stark effect by correspondence principle II171
Of Stark effect in wave mechanics III112
Internal quantum number Introduction by Sommerfeld of III181
Of half-integer (Landé) II182
Selection rule for II186
Irreversible measuring apparatus
Green’s model of III290
Irreversible process
In measuring apparatus III275, 289
Ishino II235
Ishiwara
Derivation of Balmer-like formula II113
Ishiwara
Quantum condition of II111

Jammer II88, 101, 157
Jeans
   Attempt to defend classical radiation theory II14
   → Rayleigh
Jordan
   Irreversible process and observation apparatus III289
   Quantization of de Broglie waves obeying FD statistics III225
   Speciality of spin angular momentum III218
   Transformation theory by means of probability amplitudes III184
   → Born
   → Heisenberg-
Jordan-Klein
   Quantization of charged de Broglie waves III226
Jordan-Pauli
   Quantization of free magnetic waves III227
Jordan-Wigner
   Quantization of de Broglie waves using anticommutation relation III225

Kamerlingh Onnes → Ehrenfest Kelvin
   Ether vortex model of atoms I75, 76
Mechanical model of spectral series I69
Model of atoms without nucleus
   terms of electrons I117
Kelvin model
   And property of matter I124
   Characteristics of I127
   Stability of I122
Kennard
   Commutation relation and uncertainty relation III241
Kirchhoff
   Law of black body radiation I5
Kiuti III91
Klein
Index for Volumes I, II, III

Relativization of wave equation
III202
→ Jordan-Klein-Nishina

Calculation of Compton effect by
Dirac’s electron wave equation
III232

Kopenhagen spirit → Copenhagen
spirit

Kossel
Analysis of characteristic X-rays
II130
Electron configuration of atoms
II196
→ Sommerfeld-Kramers

Calculation of intensity of Start
effect using correspondence
principle II171
Dispersion formula and induced
radiation II273
WKB approximation of wave
equation III92
→ Bohr-Kramers-Heisenberg

Generalization of dispersion formula
II282

Kuhn
Sum rule III113

Kunsman → Davisson-Kurlbaum → Rubens-Kármán → Born-

Ladenberg-Reiche
Dispersion formula and transition
probability II272

Larmor
Discussion on stability of atoms I99
General theory of Zeeman effect
based on electron theory I95
Radiation theory in terms of electron
motion I98
Larmor frequency
Of Zeeman effect I95; II142

Larmor precession
And antimagnetism I137
In the Zeeman effect I95

Lanczos
Attempt to formulate matrix
mechanics by integral equation
III57, 174, 189
Lanczos’s integral equation And the
wave equation III61
Lanczos’s unit integration kernel And
δ-function III60, 174
Landé anomalous g-factor of s terms
And the magneto-mechanical
anomaly II183
g-factor of multiplet terms II187
Half-integer quantum number and

g-factor II182
Similarity of X-ray and optical
doublet terms II209

Langevin
Larmor precession and antimagnetism
I137

Langmuir
Improvement of Lewis’s octet theory
II198
Value of chemical combination heat
of hydrogen molecule II99

Laser beams
Nagasaki’s analysis of interference of
III298

Lenard
Cathode rays and Larmor ray II100
Measurement of absorptive power for
cathode rays I164
Model of atoms in terms of
Dynamiden I163
Quantitative study of photoelectric
effect I34, 46

Lewis
Octet theory of electron
configuration II197

Light quantum
And anomalous and normal
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>scatterings of light (Smekal)</td>
<td>II279</td>
</tr>
<tr>
<td>And Compton effect</td>
<td>II236</td>
</tr>
<tr>
<td>And Doppler effect (Schrödinger)</td>
<td>II233</td>
</tr>
<tr>
<td>And photoelectric effect (Einstein)</td>
<td>I44</td>
</tr>
<tr>
<td>And recoil of atom (Einstein)</td>
<td>II151</td>
</tr>
<tr>
<td>And reflection and diffraction of X-rays (Duane, Compton)</td>
<td>II223, 241, 2442</td>
</tr>
<tr>
<td>And Stokes law (Einstein)</td>
<td>I43</td>
</tr>
<tr>
<td>de Broglie’s assumption of a minute least mass for light quantum</td>
<td>II248</td>
</tr>
<tr>
<td>Lorentz’s paradoxes</td>
<td>II34</td>
</tr>
<tr>
<td>Einstein’s presentation of the hypothesis of</td>
<td>I43</td>
</tr>
<tr>
<td>Quantization of radiation as an aggregate of</td>
<td>III206</td>
</tr>
<tr>
<td>Bose statistical mechanics</td>
<td>II255, 258</td>
</tr>
<tr>
<td>Limit of observation</td>
<td>III272, 274</td>
</tr>
<tr>
<td>Difference with the limit of cognition of</td>
<td>III272, 274</td>
</tr>
<tr>
<td>Lindemarn-Nernst</td>
<td>II22</td>
</tr>
<tr>
<td>Lodge</td>
<td></td>
</tr>
<tr>
<td>Confirmation of Zeeman effect</td>
<td>I94</td>
</tr>
<tr>
<td>Eether montistic theory of atoms</td>
<td>I76</td>
</tr>
<tr>
<td>London</td>
<td></td>
</tr>
<tr>
<td>Indication of wave function as state quantity in Hilbert space</td>
<td>III169</td>
</tr>
<tr>
<td>Wave equation and canonical transformation</td>
<td>III166</td>
</tr>
<tr>
<td>→ Heitler-Lummer-Pringsheim</td>
<td>I23, 32</td>
</tr>
<tr>
<td>Lyman</td>
<td></td>
</tr>
<tr>
<td>spectral series</td>
<td>I64</td>
</tr>
<tr>
<td>Machismic viewpoint</td>
<td></td>
</tr>
<tr>
<td>Critique by Schrödinger of</td>
<td>III100</td>
</tr>
<tr>
<td>Magneto-mechanical anomaly</td>
<td></td>
</tr>
<tr>
<td>And electron spin</td>
<td>II222</td>
</tr>
<tr>
<td>Of electron specific charge</td>
<td>II190, 214</td>
</tr>
<tr>
<td>Magneto-optical anomaly</td>
<td></td>
</tr>
<tr>
<td>Of $g$-factor of s-terms</td>
<td>II191</td>
</tr>
<tr>
<td>Magyar-Mandel</td>
<td>III296</td>
</tr>
<tr>
<td>Mandel → Magyar-</td>
<td></td>
</tr>
<tr>
<td>Pfleegor-Marden</td>
<td>II21</td>
</tr>
<tr>
<td>Margenau</td>
<td></td>
</tr>
<tr>
<td>Necessary condition of measurement</td>
<td>III289</td>
</tr>
<tr>
<td>Marsden → Geiger-Matrix element</td>
<td></td>
</tr>
<tr>
<td>And representation of $q$-number</td>
<td>III50</td>
</tr>
<tr>
<td>Representation by wave function of</td>
<td>III95, 119</td>
</tr>
<tr>
<td>Matrix mechanics</td>
<td></td>
</tr>
<tr>
<td>And angular momenta</td>
<td>III29</td>
</tr>
<tr>
<td>And anomalous Zeeman effect</td>
<td>III31</td>
</tr>
<tr>
<td>And Stark and Zeeman effects</td>
<td>III42</td>
</tr>
<tr>
<td>Application to vibration of string</td>
<td>III35</td>
</tr>
<tr>
<td>Born-Wiener’s attempt at application to aperiodic motion of</td>
<td>III61, 129</td>
</tr>
<tr>
<td>Lanczos attempt at treatment in terms of integral equation of</td>
<td>III57</td>
</tr>
<tr>
<td>Difficulty in the case of many body systems</td>
<td>III140, 147</td>
</tr>
<tr>
<td>Dirac’s generalization to aperiodic motion of</td>
<td>III173</td>
</tr>
<tr>
<td>Mathematical equivalence to wave mechanics of (Schrödinger)</td>
<td>III97, 101</td>
</tr>
<tr>
<td>Of many-body systems</td>
<td>III26</td>
</tr>
<tr>
<td>(Born-Heisenberg-Jordan)</td>
<td></td>
</tr>
<tr>
<td>Of systems of one degree of freedom</td>
<td>III19</td>
</tr>
<tr>
<td>Pauli’s application to hydrogen atoms of</td>
<td>III38</td>
</tr>
<tr>
<td>Relativization of (Dirac)</td>
<td>III203</td>
</tr>
<tr>
<td>Mayer</td>
<td></td>
</tr>
<tr>
<td>Floating magnets</td>
<td>I108</td>
</tr>
<tr>
<td>Mehra-Rechenberg</td>
<td>II88, 206</td>
</tr>
<tr>
<td>Michelson</td>
<td></td>
</tr>
</tbody>
</table>
Confirmation of Zeeman effect I94
Michelson's distribution formula
  Improvement by Wien of I14; II8
Molecular-theoretical derivation of I13; II7
Millikan-Bowen critique of
  Bohr-Coster's theory II210
Similarity of X-ray and optical
doublet terms II209
Mixture
  As statistical aggregate
    (v. Neumann) III260
Model of atom
  Bohr's II82
  Final determination of the models of atom with and without nucleus I207, 221
  Kelvin's ether voltex I75, 76
  With nucleus by Nagaoka I169
  Without nucleus by Thomson I106, 129, 140
  Without nucleus in terms of electrons by Kelvin I117
Moseley
  Verification of the Bohr model II99
Multiple scattering
  And β-ray scattering (Crowther) I207, 218
  Thomson's theory of I207, 218
Multiplet lines
  Sommerfeld's displacement law of I180
Nagaoka
  Presentation of model of atom with nucleus I169
  Reply to Schott's critique I180
Nagaoka model
  And ideal atoms I181
  And line- and band-spectra I169, 177
  And radioactivity I165, 175
  And intensity of spectral lines I197, 206
Problem of stability of I180
Quantity of central charge in I171
Quantity of central mass in I181
Nagasaki
  Analysis of interference of two laser beams III298
Natural radiation
  Planck's I110
Negation of causality
  Bohr's support for III249
  On the basis of uncertainty relation (Heisenberg) III247
Nernst → Lindemann
Nicholson
  Indication of quantum nature of angular momentum I177
  Radius of an atom of the Nagaoka type I172, 73
  Theory of corona spectra I167
Nishina → Klein-
  Nonmechanical compulsion to lighting electrons (Bohr) II209, 216; III103
  Negation by Pauli of II216
Nordheim → Hilbert–Nuclear spin
  And symmetry of wave function (Heisenberg) III221
  Dennison's experimental value for hydrogen atom of III222
Number of electrons
  Drude's determination about atoms of I191
  Determination about hydrogen atom of (Bohr) I229, II80
  Investigation about atoms of (Thomson) I182, 193
Number representation
  Of the wave equation (Dirac) III209
Objective interaction
  As origin of statistical nature of quantum mechanics III271
Observation apparatus
  Boundary between object of observation and III261, 274, 275
Conditions for v. Neumann type of III263, 287, 292
Green's model of III290
Of the so-called negative experiment type III293
Pauli's model of III277
Observation in quantum mechanics
Necessary condition for III276, 288
Observation theory
On the basis of principle of psycho-physical parallelism (v. Neumann) III261
Octet theory
About electron configurations by Lewis II197
Improvement by Langmuir of II198
Ortho-helium II201
Out-radiation
Similarity to radioactivity II149
Par-helium II201
Paschen → Ritz–Paschen–Back effect II183, 218, 222
Pauli
Conjugate configuration of electrons and reciprocal theorem II217
Indication of the doubleness of electron II215
Introduction of spin wave function III216
Matrix mechanics of hydrogen atom III38
Matrix mechanics of Stark and Zeeman effects III42
Model of measuring apparatus III277
Negation of Bohr's idea of nonmechanical compulsion II216
Presentation of the exclusion principle (the general rule) II217, 219
Quantization of direction and magnetization of gas II192
Relativistic correction of g-factor II213
Wave function as probability amplitude III180
→ Heisenberg–
→ Jordan
Perihelion
Of relativistic electron orbit II125
Person
Magnetic force between self-rotating electrons II220
Perturbational method
In matrix mechanics III28
In wave mechanics (time-dependent) III158, 163
In wave mechanics (time-independent) III107
Pfleegor-Mandel III296
Pfund
Spectral series I63
Phase of wave
And Hamilton's principal function III84
Phase velocity
Of the de Broglie wave II250, III70
Phase wave (virtual wave) de Broglie's II250, III85
Photoelectric effect
And Stokes law I43
And the hypothesis of light quantum I44
Derivation by wave mechanics (Wentzel) II200
Quantitative study of (Lenard) I34, 46
Physical mechanics
Name of wave mechanics (Schrödinger) III93
Pickerling II89
Planck
Complexion number I28; II15
Distribution formula of the black body radiation I25; II11
Hypothesis of elementary action II47
Hypothesis of natural radiation II10
Index for Volumes I, II, III

Quantum condition II136
Resonator I17, 24; II9
Planck constant → Elementary action
Planck’s distribution formula
And the entropy of resonators I25, 39; III11, 15
Derivation by Bose of II255
Derivation by Debye of II18, 261
Derivation by Einstein of II150, 258
Two limits of I37, 49
Podolsky → Einstein-
Poincare

Doubts on quantum condition II145
Poisson bracket equation
And the commutation relation III47
Positive electron
Dirac’s theory of III234
Principle of psycho-physical parallelism
In the Copenhagen spirit III252
Pringsheim → Lummer-
Probability amplitude
As interpretation of wave function (Pauli) III183
Jordan’s transformation theory by means of III184

$q$-algebraic treatment
Of angular momentum III53
Of equation of electron orbit III50

$q$-number
Dirac’s III47
Matrix element and representation of III50

Quantization of direction
And Stern-Gerlach’s experiment II192
And magnetization of gas II192
Sommerfeld’s introduction of II123

Quantum condition
Analogy to Frenel’s band theory of III89
And adiabatic hypothesis II165
And Balmer’s formula II118

And reflection and diffraction of X-rays II241-243
And the WKB approximation III92
Rubinowicz’s application to cavity radiation of II160
Sommerfeld, Epstein, Schwarzschild’s application to the Stark effect of II121, 132, 136
Debye and Sommerfeld’s application to Zeeman effect of III141, 142
Difficulty for aperiodic motions of III195
Doubt’s by Poincare on II145
Equality of Sommerfeld’s and Planck’s III140
Intensified (Born-Jordan) III23
Brillouin’s interpretation by inneratomic elastic wave of II244
de Broglie’s interpretation by virtual (phase) wave of II249
Ishiwara’s III111
Limitation of II128
Planck’s III136
Quantum-theoretical (Heisenberg) III11, 14
Schrödinger’s replacement with a variation principle of III74, 77
Sommerfeld’s III116
Viewpoints from energetics and atomism of II146
Wilson’s III108

Quantum number
As premise of the quantum condition II116
Bohr’s assumption of constancy of II218
Schrödinger’ derivation from wave equation of III76
Quantum-theoretical multiplication formula
Heisenberg’s III110
Quantum-theoretical quantum condition
Heisenberg’s III10, 14

Ramusaner II260

Rayleigh

Difficulty in classical theory of spectra I73

Dynamical model of spectral series I66

Rechenberg → Mehr-

Reciprocal theorem

Of black body radiation (Kirchhoff) I6

Of the conjugate electron configuration (Pauli) II217

Reiche → Ladenberg-

Relativistic

Fine structure II127; III202, 232

Matrix mechanics III203

Wave equation III150, 154, 203, 231

Resonator

Entropy of I19, 25, 31; II15

Radiation density and average energy of I30, 37; II6, 10

Reyleigh-Jeans

Distribution formula of black body radiation 123; II12

Reyleigh-Jeans’s distribution formula

And the equipartition principle of energy I37; II12

As limit of Planck’s formula I38

Richardson II190

Ritz atomic magnet model of spectral series II60

Combination rule of spectral terms II64

Plate vibration model of spectral series II56

Ritz-Paschen

Spectral series I64

Robertson

Commutation relation and the uncertainty relation III241

Rosen → Einstein-Rosenfeld II101

Rubens-Kurlbaum I24, 312

Rubinowicz II101

Application of quantum condition to cavity radiation II160

Conservation of angular momenta and selection rule II159

Runge

Rule for anomalous Zeeman effect II181

→ Sommerfeld-

Rutherford

Final determination of model of atom with or without nucleus I205, 221

Rutherford scattering formula Gordon’s derivation by wave mechanics of III200

Wentzel’s Born approximation calculations of III199

Rutherford-Anderade II156

Rydberg

Electron configuration in rare gas atoms II198

Spectral series I64, 65

Schott

Instability of Nagaoka model I180

Intensity of radiation from electron rings I197

Oscillation of electron rings I170

Schrödinger

Assumption of $\psi (d\bar{\psi}/dt)$ as charge density III110, 120

Assumption of $\psi \bar{\psi}$ as charge density III120

Assumption of $\psi \bar{\psi}$ as weight function III121

Attempt to interpret wave function as catalog of expectation III269

Critique of Machismic viewpoint III101

Derivation of Balmer’s formula III76

Dispersion formula in wave mechanics III118

Example of a cat (critique of
observation theories of v. Neumann-type) III268
Light quantum and Doppler effect II231
Mathematical equality of matrix and wave mechanics III96, 101
Objective interaction as origin of statistical nature of quantum mechanics III271
Relative intensity of Stark effect III112
Relativization of wave equation III155
Replacement of quantum condition using variation principle III74, 77
Theory of gases by means of de Broglie waves III64
Time-dependent wave equation III115, 155, 181
Time-independent wave equation III77
Schuster I78
Schwartzschild
Quantum condition and Stark effect II136
Second quantization
Of Dirac electron wave equation III228, 234
Quantization of de Broglie wave field III208, 225, 226
Selection rule
And internal quantum number II186
And conservation of angular momentum II159
Derivation by matrix mechanics of III30
Derivation from correspondence principle II163
Sommerfeld's provisional II120, 129
Shimizu II221
Siegbahn
Measurement of characteristic X-rays II130
Signal velocity → Group velocity
Simon → Compton-
Slater Idea of virtual radiation II262
→ Bohr-
Smekal
Aperiodicity of translational motion in magnetic field II195
Light quanta and normal and anomalous scatterings of light II278
Sommerfeld
Derivation of formula for spectral terms II179
Hypothesis of action integral II48
Introduction of internal quantum number II181
Introduction of quantum condition II116
Magneto-optical anomaly of g-factor of s-terms II191
Permanence of multiplicity II180, 186
Provisional selection rule II120, 129
Quantum condition and Balmer's formula II118
Quantum condition and quantization of direction II123
Quantum condition and Stark effect II121
Quantum condition and Zeeman effect II142
Theory of relativistic fine structure II127
Sommerfeld-Heisenberg
Selection rule for internal quantum number II186
Sommerfeld-Kossel
Displacement law of spectral series II180
Sommerfeld-Runge
Wave equation of light of short
wavelengths III83
Specific charge
  Measurement by Thomson of
cathode-ray electron I100, 105
  Measurement by Wien of anode-ray
particle II96
  Measurement by Zeeman of electron
I94
Specific heat of solids Born-Kármán’s
theory of II24
  Debye’s theory of II23
  Einstein’s theory of II19, 21
Spectral term formula
  Ritz’s II63
Sommerfeld’s II179
  Balmer’s I63
  In the Bohr model II88
  Kayser-Ringe’s I65
  Rydberg’s I64, 65
Spectrum theory
  Breakdown for oscillation of electron
rings of I206
  Difficulty in the classical theory of
I74
  Nicholson’s II67
  Ritz’s II56, 60
Spin
  And magneto-mechanical anomaly
II221
  Darwin’s attempt to introduce
vector-type III214
  Pauli’s introduction of wave function
for III215
  Specialty of spin angular momentum
(Jordan) III218
  Uhlenbeck–Goudsmidt’s introduction
of II219
Spin-orbit interaction
  In Dirac electron wave equation
III232
  Thomas-Frenkel’s II223
Stark effect
  Pauli’s application of matrix
  mechanics to III42
Sommerfeld, Epstein and
  Schwartzschild’s application of
quantum condition to II121, 132,
136
  Wentzel’s and Schrödinger’s
  application of wave mechanics to
III92, 108
Relative intensity by correspondence
  principle (Kramers) in II171
Relative intensity by wave mechanics
(Schrödinger) in III112
Bohr’s treatment by a
  pre-correspondence-principle of
II143
State
  As an essentialistic concept III273
State probability
  And matrix element of canonical
  transformation III172
  Introduction by Born of III129, 162
State vector
  And unified formulation of matrix
  and wave mechanics III193, 197
Stationary state
  Assumption in the Bohr model of
II86
Statistical aggregation
  Introduction by v. Neumann of
III257
Statistical nature (in quantum
  mechanics)
  Making much in the Copenhagen
  spirit of III252
  Objective interaction as origin of
III271
Statistical operator
  Introduction by v. Neumann of
III258
Stefan-Boltzman
  Law of the black body radiation I7
Stereo-structural logic
  Of quantum mechanics (Taketani)
Index for Volumes I, II, III

III272, 275
Stern-Gerlach’s experiment
And the anomalous Zeeman effect II195
Difficulty of interpretation of II192
Stewart II190
Stieltjies’s integral-in v. Neumann’s formulation III195
Stokes-Law of photoelectric effect I44
Stoner
Revision of Bohr’s electron configuration II211
Sucksmith-Bates II191
Sugiura-calculation of covalent bonds III223
Sum rule
Kuhn’s and Thomas’s III13
Sutherland II21

Taketani
Conflict between field and matter II225; III235
Critique of Bohr’s complementarity III253
Irreversible process and the objective boundary in observation III275
Stereo-structural logic in quantum mechanics III271, 275
Tamm
Calculation of Compton effect by quantum field theory III233
Theory of electron
Lorentz’s I77
Thomas, L. H.
Derivation of spin-orbit interaction II222
Thomas, W.
Sum rule III14
Thomson
Analogy of floating magnets and periodic law I108
Determination of Avogadro’s number (by method of electrolysis) I112
Discussion of number of inneratomic electrons I182, 186, 190, 193
Equality of positive and negative elementary charges I114
Measurement of elementary electric charge (by method of gas ionization) I111
Measurement of specific charge of free cathode ray electron I100, 104
Model of atoms without nucleus by corpuscles I106 129, 141
Specific charge of anode ray particle and mass of positively charged particle I196
Theory of magnetic body I128, 136
Theory of multiple scattering I207, 218
Thomson model
And electron temperature I159
And radioactivity I161
And construction of periodic system I150
And electron mass as origin of atomic mass I158
And intensity of radiation I135
And intensity of spectral lines I198, 206
And oscillation of inneratomic electrons I154
And valence I156
Stability condition and oscillation modes of I141, 150
Townsend
Equality of positive and negative elementary charges I114
Measurement of elementary charge (by method of ion diffusion) I113
Transformation function
And Born’s scattering probability (Dirac) III182
Introduciton by Dirac of III176
Wave function as III179
Transformation theory
  Dirac's, Jordan's III179, 189
Transition probability
  And dispersion formula II272, 274
  And quantum theoretical amplitude
  III17
  And correspondence principle II170
Introduction by Einstein of II149
  Of absorption and induced emission
  of light in wave mechanics III158, 164
  Of spontaneous emission of light in
  quantum mechanics III212
Reversibility of (Born) III165
Uhlenbek-Goudsmit
  Introduction of electron spin II219
Uncertainty principle
  Arbitrary interpretation in
  philosophical discussions of III244
  Name of the uncertainty relation
  III244
Uncertainty relation
  Heisenberg's derivation by means of
  thought experiments III240
  Derivation from commutation
  relation of III242
  Negation of causality on the basis of
  (Heisenberg) III247
Unification
  Framework of III164
  Moment of III180
Unified formulation of matrix and wave
  mechanics III193, 197
Uniform aggregation
  As a statistical aggregation
  (v. Neumann) III259
Uniformising variable
  for multi-periodic system III49
Unit operator
  Splitting in v. Neumann's theory of
  III195
Viewpoint of atomism
  About the quantum condition II146
  About thermal phenomena I6, 17
Viewpoint of energetics
  About quantum condition II146
  About thermal phenomena I6, 17
Villard II156
Virtual orbit
  Bleit's II275
Virtual oscillator
  Bohr-Kramers-Slator's II267, 275
Virtual radiation field Slator's II265
v. Neumann
  Mixed aggregation as a statistical
  aggregate III260
  Statistical aggregation III257
  Statistical operator III259
  Theory of observation based on
  principle of psycho-physical
  parallelism III261
  Unified formulation of matrix and
  wave mechanics III193, 197
  Uniform aggregation as a statistical
  aggregate III259
  → Hilbert-
  v. Neumann's type of measurement
  apparatus Conditions for III263,
  287, 292
Voigt
  Model of Paschen–Back effect of
  Na-D line II183
Vortex atom
  Kelvin's I75, 76
Wave equation
  Number representation of (Dirac)
  III209
  Dirac's relativistic electron III231
  Dirac, Schrödinger, Klein, Fock,
  Gordon's relativization of
  III150, 155, 202
  Schrödinger's and Dirac's
  time-dependent III116, 154, 181
Schrödinger's time-independent
Wave function
As probability amplitude (Pauli) III183
As a state quantity in Hilbert space (London) III169
As a transformation function (Dirac) III179
As expectation catalog (Schrödinger) III269
As probability amplitude transformation in theory (Jordan) III184
Attempt at interpretation as ghost field of III126
Attempt at interpretation as guiding wave of II251, III126
Born's assumption of $\psi \bar{\psi}$ as probability of existence III129, 162
Of spin (Pauli) III216
Permutation property and nuclear spins III221
Permutation property and spectral terms III219
permutation property for a many-body system III144, 149
Schrödinger's assumption of $\psi \bar{\psi}$ weight function III121
Schrödinger's assumption of $\psi \bar{\psi} \propto$ density of electric charge III118
Schrödinger's assumption of $\psi \frac{d\bar{\psi}}{dt} \propto$ density of electric charge III110, 118
Schrödinger's representation of matrix element by III95, 119

Wave mechanics
Applied to diatomic molecules III88
Applied to hydrogen atom III74
Applied to Planck's oscillator III88
Born's application to scattering problem of III131

Born's introduction of particle nature to III129, 135
Mathematical equality to matrix mechanics of (Schrödinger) III97, 101
Weight function
As an interpretation of $\psi \bar{\psi}$ (Schrödinger) III121

Weiss
Molecular magnetic field of gases II60

Wentzel
Born approximation of Rutherford scattering formula III199
Regular and irregular doublet X-ray terms II205
Wave equation and photoelectric effect III200
WKB approximation for Stark effect III91
WKB approximation for wave equation III90, 92

Weyl
Commutation relation and uncertainty relation III241
Whole and parts in quantum mechanics III273

Whiddington
Discussion of velocity loss of charged particles I222
Energy of generating characteristic X-rays II81, 98

Whole and parts
In quantum mechanics (Weyl) III273

Wien
Displacement law of black body I8; II8
Distribution formula of black body radiation I16; II8
Improvement of Michelson's distribution formula II8
Measurement of specific charge of
anode rays 1196
Wiener → Born-Wien's formula
And hypothesis of light quanta 143
As a limit of Planck's formula 149
Wigner → Jordan-Wilson, H. A.
Development of cloud chamber 1110, 111
Measurement of elementary charge
(by method of falling drop) 1113
Wilson, W.
Quantum condition 1108
WKB approximation
Applied to Stark effect 11891
For wave equation 1186, 91
Wright 1171
X-ray doublet terms
Regular and irregular 11205
Screening and relativistic 11205
Similarity to optical doublet terms of 11209
Zeeman
Discovery of Zeeman effect 1185
Measurement of specific charge of electron 1194
Zeeman effect
And Larmor frequency 1196; 11142
And verification of Lorentz electron theory 1191
Bohr's pre-correspondence-principal treatment of 11143
Debye and Sommerfeld's application of quantum condition to 11141, 142
Pauli's application of matrix mechanics to 11142
Relative intensity by matrix mechanics of 11130