

Wolfgang Pauli

General Principles of Quantum Mechanics



Springer-Verlag Berlin Heidelberg New York

Wolfgang Pauli

General Principles of Quantum Mechanics

**Translated by
P. Achuthan and K. Venkatesan**



**Springer-Verlag
Berlin Heidelberg New York 1980**

**The original German edition of this work was published under the title:
Handbuch der Physik, Vol. 5, Part 1: Prinzipien der Quantentheorie I, 1958**

**ISBN 3-540-02289-9 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-02289-9 Springer-Verlag New York Heidelberg Berlin**

**The tenth chapter of this book is translated from Part B,
Sections 6–8 of Pauli's article in: Handbuch der Physik, Vol. 24, Part 1,
1933, edited by H. Geiger and K. Scheel**

**ISBN 3-540-09842-9 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-09842-9 Springer-Verlag New York Heidelberg Berlin**

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Under § 54 of the German Copyright Law where copies are made for other than private use, a fee is payable to the publisher, the amount of the fee to be determined by agreement with the publisher.

**© by Springer-Verlag Berlin Heidelberg 1980
Printed in India**

**Typesetting and Printing: Allied Publishers Private Ltd., New Delhi, India.
Bookbinding: Graphischer Betrieb, K. Tritsch, Würzburg.
2153/3140-543210**

Pauli and the Development of Quantum Theory¹

The "Relativitätstheorie" written by the 21 year old Pauli in 1921 and the "Wellenmechanik" of 1933 were both re-edited in 1958, the last year of Pauli's life. This rare distinction shows that both of these reviews have become classics of the 20th century theoretical physics literature. It is perhaps surprising that in his later years, Pauli the great critic himself had come to admit that he really was a classicist and not the revolutionary innovator he had thought himself to be in his youth.

The success story of these two review articles makes one forget that good luck had also played a part: their timing was right; they were written at a moment when the respective fields had matured towards logically complete theories. Pauli did not always have this luck. In fact, he had written two other review articles on quantum theory around 1925. But all the qualities of Pauli's writing – his logical insight, his precise formulation, his cautious judgement and his care with details – could not prevent these two articles to be obsolete when they appeared in print. The more important of them, the "Quantentheorie" of 1926, is a review of the old Bohr-Sommerfeld theory, updated only by occasional footnotes to include electron spin and the exclusion principle. For this reason and because it had also appeared in Geiger and Scheel's *Handbuch*, Pauli later called this review half jokingly the "Old Testament": It is a wealth of historic truths built on the Old Rules, but salvation could only come from the "New Testament", the "Wellenmechanik" of 1933.

Speaking of reviews one must of course not forget that Pauli has marked both the old and the new quantum theory by the stamp of his own research and his profound critical analyses. His first contribution to the old theory was his doctoral thesis on the hydrogen molecule ion written in Sommerfeld's institute and submitted to Munich University in 1921. More important was his investigation of the anomalous Zeeman effect since it culminated in the paper on the exclusion principle (January 1925) which won him the Nobel Prize of 1945.

When the foundation of the new theory was laid by Heisenberg in the summer of 1925, Pauli surprised him only a few months later by the solution of the hydrogen atom in the new matrix mechanics. And when Schrödinger the following year published his first communication on quantisation as eigenvalue problem it was again Pauli who in a letter to Jordan showed its equivalence with the matrix mechanics of the Göttingen School. Schrödinger, however, was quicker so that Pauli's proof was not published.

In 1927, after his faith in electron spin had at last been consolidated by Thomas'

¹ For more details see C.P. Enz, W. Pauli's Scientific Work, in *The Physicist's Conception of Nature*, ed. J. Mehra (Reidel, Dordrecht-Holland, 1973).

correct derivation of spin-orbit coupling, Pauli developed his non-relativistic theory of the magnetic electron, fully aware that the final theory would have to be relativistic. This was the triumph of Dirac's paper published the following year. Dirac's theory not only generalised the Pauli spin matrices but also led the way out of the dilemma of the Schrödinger-Klein-Gordon equation of 1926, which as a one-body spin-zero theory could not exclude negative probabilities. The description of this dilemma and its resolution by Dirac's multi-component wave function is one of the showpieces of Pauli's exposition in the "Wellenmechanik" of 1933. However, furtheron in the review, Pauli rejects Dirac's re-interpretation of the unoccupied negative-energy states because at the time of writing the only known positive particle was the proton. This was bad luck since by the time the article appeared in print, Anderson's discovery of the positron (1932) had already brilliantly vindicated Dirac's interpretation. For this reason, Pauli modified and considerably shortened both the introduction to the relativistic one-body problem and the section on negative energy states in the edition of 1958.

There are two other, quite innocent-looking, changes in this part of the 1958 edition. The first is a modification of the footnote concerning the tensor quantities formed with products of the γ -matrices. While in the old version Pauli expressed his dissatisfaction with the particular representation used to derive the quadratic identities of this footnote, the new version quotes Pauli's two papers of 1935 and 1936 in which the representation theory of the Dirac matrices had been fully elucidated. However, his fundamental theorem, namely that the 4×4 matrix representation is irreducible, has been mentioned already in the 1933 edition.

The second innocently looking change concerns Weyl's two-component equation for massless spin- $\frac{1}{2}$ particles which Pauli had rejected in the 1933 edition because of its parity violation. In the 1958 edition a footnote on the neutrino and parity-violating weak interactions was added, behind which are hidden the triumph and surprises of another brainchild of Pauli: the neutrino, conceived in 1930, experimentally verified in 1956 and revealed to be left-handed in the following year.

Pauli, with Dirac, Jordan and Heisenberg, had been one of the founding fathers of quantum electrodynamics and, more generally, of quantum field theory. This was in 1928 when he joined the Swiss Federal Institute of Technology (ETH) in Zurich. It was therefore natural that the 1933 edition contained a section on quantum electrodynamics. One year later, Pauli had shown (in a paper with Weisskopf) that in a second-quantised form the spin-zero Klein-Gordon equation was as satisfactory as the spin- $\frac{1}{2}$ Dirac theory. At the same time, however, this paper also showed that Dirac's argument for choosing a multi-component wave function had to be revised. This profound analysis eventually led Pauli to his famous spin-statistics theorem of 1940.

By 1958 the development of quantum electrodynamics had completely outgrown Pauli's review of 1933 so that the reprint of these sections was not justified anymore. However, these deleted sections contain two points of great concern to Pauli. The first is the question of the concept of the electric field strength and of the atomicity of the electric charge which Pauli had raised for the first time in his third published paper at the age of 19 and which culminates in the question concerning the origin of the value 1/137 of the fine structure constant.

The second point concerns the problem of the zero-point energy which Pauli used

to discuss with Otto Stern during his Hamburg years (1923–1928). While for material oscillators Stern argued convincingly in favour of the reality of the zero-point energy, for the radiation field Pauli had good arguments against its reality: the gravitational effect of the zero-point energy of radiation is zero, as is that of the filled negative-energy states in Dirac's theory of the electron. This was the first occurrence, although in a trivial form, of the problem of renormalisation of a divergent quantity.

These details may serve to show that the sections which are deleted from the 1958 edition are of considerable historical interest, at least with respect to Pauli's scientific work.¹ It is therefore a happy decision of the translators to include them as the last chapter in the present English edition. Having myself read the proofs of the 1958 edition when I was Pauli's assistant at ETH, it is a particular pleasure for me to introduce this English translation – which I have again proofread, but this time with the aim of preserving the rigour and spirit of Pauli's historic work. This translation is in fact a project which I had hoped to see realised since many years, convinced that this second classic of Pauli would be acclaimed by the English speaking scientific community. Its coming is none too soon, considering that Pauli's Collected Scientific Papers have already been translated into Russian several years ago.

CHARLES P. ENZ

Geneva, 24 February 1977

Foreword

I am very happy to accept the translators' invitation to write a few lines of introduction to this book.

Of course, there is little need to explain the author. Pauli's first famous work, his article on the theory of relativity in the *Encyclopädie der Mathematischen Wissenschaften* was written at the age of twenty. He afterwards took part in the development of atomic physics from the still essentially classical picture of Bohr's early work to the true quantum mechanics. Thereafter, some of his work concerned the treatment of problems in the framework of the new theory, especially his paper on the hydrogen atom following the matrix method without recourse to Schrödinger's analytic form of the theory. His greatest achievement, the exclusion principle, generally known today under his own name as the Pauli principle, that governs the quantum theory of all problems including more than one electron, preceded the basic work of Heisenberg and Schrödinger, and brought him the Nobel prize. It includes the mathematical treatment of the spin by means of the now so well-known Pauli matrices. In 1929, in a paper with Heisenberg, he laid the foundation of quantum electrodynamics and, in doing so, to the whole theory of quantized wave fields which was to become the *via regia* of access to elementary particle physics, since here for the first time processes of generation and annihilation of particles could be described for the case of the photons. Later on, he solved the riddle of the continuous β -spectrum, at the time seemingly violating the mechanical conservation laws, by postulating the existence of the neutrino, thus preparing the way for Fermi's detailed theory of this phenomenon and all theories of weak interactions up to the present date.

Pauli's article which is here submitted in English translation was first published in the Geiger-Scheel Handbuch der Physik in 1933. Later on, it has at my instance been reproduced with only a few very slight alterations and adaptations – and a very few omissions towards the end – in the Encyclopedia of Physics in 1958, a few months only before Pauli's demise. Thus the text of the present translation is almost the same as had first been published in German 44 years ago, about seven years after the final discovery of the true form of quantum mechanics. During the long years that have since elapsed so much progress in physics has been achieved and so many shifts in interest have happened that it may, at a first glance, seem rather amazing that there still remains enough interest for this article to justify its translation.

One of the reasons certainly lies in Pauli's having been one of the great masters of theoretical physics. But there have been several of his contemporaries whose names hold as much renown and whose original work has contributed as much as

Pauli's to the achievements of these greatest times of modern physics, but whose work has by no means kept the same popularity among the present generation. It may be added that, twenty years ago, when I asked Pauli to republish his old article in the then forthcoming Encyclopedia, it seemed not at all sure that it would meet with enough interest. What then has changed in the meantime?

The next generation – in fact, my own – who started work after the discovery of quantum mechanics was faced with the tremendous task of applying this theory to the vast realm of atoms, molecules, and later on to the world of nuclei. We just took the theory for granted, not having participated in the long and weary struggle to achieve it. To the preceding generation, a large part of whose lives had been devoted to this struggle, the new principles found and their strange deviation from classical thought had been much more essential than the broad stream of detailed application following its setting up. For the younger people, concerned with their own work, mathematical techniques, group theoretical systematization and approximations to more and more complicated problems filled the foreground of the stage. The principles, by no means forgotten, were however no longer of paramount interest and therefore banished to the background. Most of the later books, and even excellent books, on quantum mechanics have thus only in a comparatively brief way dealt with the underlying general ideas. Pauli's article, that does, is therefore not just another book among the large number already existing on the same subject, but stands quite in its own right.

During the last ten years or so there seems to be a fast growing feeling that we have struck another barrier to progress in the region of very high energies. We have become more conscious of our still imperfect understanding of the principles underlying the structure of elementary particles. This may well mean reconsidering our general position in the years to come to find out where and why current theories may become inadequate. The situation is thus – on another level, of course – not so very dissimilar to the one before the big breakthrough of 1925 to quantum mechanics as we know it today. The principles of this now so well established theory as represented in Pauli's article are thus gaining new interest and are becoming increasingly important.

2 August 1977

S. FLÜGGE

Translators' Preface

The subject matter of the first nine chapters of the present book is a translation of the article "Die allgemeinen Prinzipien der Wellenmechanik" (which we have called "General Principles of Quantum Mechanics" in keeping with modern usage) by Wolfgang Pauli (Nobel Laureate) which appeared in Vol. V, Part I of *Handbuch der Physik* (1958) edited by S. Flügge. The tenth chapter is translated from Part B, Sections 6–8 of Pauli's article in *Handbuch der Physik* (1933) Vol. 24, Part I, edited by Geiger and Scheel. The usefulness at the present time of a translation of an article originally written in 1933 may appear questionable for one who may not be aware of Pauli's way of presenting physics. The fact that Pauli's article of the 1933 edition was the only one which was reproduced practically in its original form, in the 1958 edition, speaks for the depth and clarity of his treatment which delved directly into the essentials of the subject, brushing aside minor details. We need not be surprised by this because Pauli had already written a remarkable book on the Theory of Relativity when he was just a student which is even now one of the best books on Relativity. It is this critical approach of Pauli to any subject which entitles him to be called "conscience keeper of physics" or "living conscience of theoretical physics".

Since the original work was in the form of an article, we had to make certain alterations in arrangement of the matter. Instead of the division into Part A and Part B for the non-relativistic and relativistic sections respectively, the 27 sections have been grouped into ten chapters. This rearrangement may not be the ideal one but over-riding considerations of keeping the numbering of equations and order of presentation intact led to the present form. Some lengthy sections have been divided into sub-sections. We have added a few remarks here and there as footnotes and apart from additional references, these are distinguished from those of the author by asterisks. In the various sections, the author's footnotes have been renumbered continuously.

A major addition has been made to the article of the 1958 edition. The sections 6, 7 and 8 of the article in the 1933 edition had been dropped in the 1958 edition, since a more detailed article on Quantum Electrodynamics by Källén (now brought out as a book) was included in the same volume. We have restored now these three sections for completeness. The first two sections (6 and 7), as representing the early development of quantum electrodynamics in the hands of Heisenberg, Pauli and others, are still useful.

Some minor errors have been corrected, but we shall be thankful if the readers coming across any others can kindly communicate them to us.

This book will be of immense value to a serious student who wishes to attain a

critical knowledge of quantum mechanics. Some special features of the book are the fine presentation of the theory of measurements, discussion of transformation groups, study of the behaviour of eigenfunctions of identical particles, application of semiclassical theory of radiation to coherent properties of radiation, the careful handling of the non-relativistic limit of the Dirac equation and the WKB approximation for Dirac equation.

A word about the notation: We have preserved the author's notation of representing operators by bold letters, vectors by arrows over them, scalar products of vectors by enclosing them in parenthesis and vector products with square brackets.

We are indeed fortunate to get a foreword for the book from Prof. Dr. S. Flügge of the University of Freiburg and editor of the Handbuch der Physik Series and a write-up on "Pauli and the Development of Quantum Theory" by Prof. Dr. C.P. Enz of the University of Geneva. We very sincerely thank them both for providing these forewords.

To Professor Enz we owe a special debt of gratitude for a critical reading of the manuscript and for suggesting valuable corrections which have been incorporated in the translation.

We express our gratefulness to Professor S. K. Srinivasan for his kind encouragement throughout this work. It is a pleasure to thank Dr. R. Subramanian for reading the draft and suggesting improvements on the style of presentation. Mr. K. V. Venkateswaran did a good job in typing the manuscript efficiently. We wish to record with thanks the splendid co-operation of Mr. N. K. Mehra and the Springer Verlag.

Department of Mathematics
Indian Institute of Technology
Madras 600036
October 1977

P. ACHUTHAN
K. VENKATESAN

Contents

CHAPTER I

The Uncertainty Principle and Complementarity

1.	The Uncertainty Principle and Complementarity	1
2.	The Measurement of Position and Momentum	8

CHAPTER II

Schrödinger Equation and Operator Calculus

3.	The Wave Function of Free Particles	13
4.	The Wave Function of a Particle Acted on by Forces	23
5.	Many-Particle Interactions – Operator Calculus	31

CHAPTER III

Stationary States and the Eigenvalue Problem

6.	Stationary States as Eigenvalue Problem	42
----	---	----

CHAPTER IV

Matrix Mechanics

7.	General Transformations of Operators and Matrices	54
8.	The General Form of the Laws of Motion	61

CHAPTER V

Theory of Measurements

9.	Determination of the Stationary States of a System through Measurement: General Discussion of the Concept of Measurement	67
----	---	----

CHAPTER VI

Approximation Methods

10.	The General Formalism of Perturbation Theory	79
(a)	Stationary State Perturbation Theory	79
(b)	Time-dependent Perturbation Theory	83
11.	Adiabatic and Sudden Perturbations	86

(a) Adiabatic and Sudden Perturbations of a System	86
(b) The Most General Statement on Probability in Quantum Mechanics	89
12. The W K B Approximation	91
(a) Limiting Transition to Classical Mechanics	91
(b) Relation to the Old Quantum Theory	98
 CHAPTER VII	
Identical Particles, Spin and Exclusion Principle	
	103
13. Hamiltonian Functions with Transformation Groups. Angular Momentum and Spin	103
(a) Group Theoretical Considerations	103
(b) Wave Functions for Particles with Spin	112
14. The Behaviour of Eigenfunctions of Many Identical Particles Under Permutation. The Exclusion Principle	116
(a) Permutations and the Symmetry of Eigenfunctions	116
(b) The Exclusion Principle	119
 CHAPTER VIII	
Semiclassical Theory of Radiation	
	130
15. Treatment of the Radiation Processes Based on the Correspondence Principle	130
16. Application of the Semiclassical Theory to the Coherence Properties of Radiation	140
 CHAPTER IX	
The Relativistic One-Particle Problem	
	145
17. Introduction	145
18. Dirac's Wave Equation for the Electron. Free Particle Case	145
19. Relativistic Invariance	150
20. The Behaviour of Wave-Packets in the Free Particle Case	158
21. The Wave Equation when Forces are Present	163
22. Approximations of the Dirac Theory: The Non-relativistic Quantum Mechanics of Spin as First Approximation	168
23. Approximations of the Dirac Theory: Limiting Transition to the Classical, Relativistic Particle Mechanics	172
24. Transitions to States of Negative Energy. Limitations of the Dirac Theory	174
 CHAPTER X	
Quantum Electrodynamics	
	177
25. Quantisation of the Free Radiation Field	177
(a) Classical Theory	177
(b) Quantisation	180
(c) Limits of Accuracy on the Measurement of Field Strengths	187
(d) Transition to the Configuration Space of Photons	189
26. Interaction Between Radiation and Matter	192
27. Self-Energy of the Electron. Limits of the Present Theory	201
 Bibliography	205
Index	209

CHAPTER I

The Uncertainty Principle and Complementarity¹

The last decisive turning point of quantum theory came with de Broglie's hypothesis of matter waves,² Heisenberg's discovery of matrix mechanics,³ and Schrödinger's wave equation,⁴ the last establishing the relationship between the first two sets of ideas.* With Heisenberg's uncertainty principle⁵ and Bohr's fundamental discussions⁶ thereon the initial phase of development of the theory came to a preliminary end.

The foundations of the theory relate directly to the problems of wave-particle duality of light and matter. The theory leads to the long-sought-for solution of the problem and gives a complete and correct description of the connected phenomena.† This solution is obtained at the cost of abandoning the possibility of treating physical phenomena objectively, i.e. by abandoning the classical space-time and causal description of nature which essentially rests upon our ability to separate uniquely the observer and the observed.

As an illustration of the difficulties which beset the simultaneous use of the photon and wave concepts of light, let us consider an approximately monochromatic point source of light, placed before a diffraction grating. For the sake of simplicity, assume that its resolving power is infinitely large. According to the wave theory, the light diffracted by the grating can pass to only certain definite points corresponding to

* W. Heisenberg, *The Physical Principles of the Quantum Theory*, Dover Publications, Inc., New York (1949); N. Bohr, *Atomic Theory and the Description of Nature* (in the following cited as A.a.N.), Cambridge Univ. Press (1961); Solvay-Congress (1927); L. de Broglie, *Introduction à l'étude de la mécanique ondulatoire*, Paris (1930); E. Schrödinger, *Lectures on Wave Mechanics* (in German), Berlin (1928).

† L. de Broglie, Ann. d. Phys. (10) 3, 22 (1925). (Thesis, Paris, 1924); cf. also A. Einstein, Berl. Ber. (1925), p. 9.

² W. Heisenberg, Z. Physik 33, 879 (1925); cf. also M. Born and P. Jordan, Z. Physik 34, 858 (1925); M. Born, W. Heisenberg and P. Jordan, Z. Physik 35, 557 (1926); P.A.M. Dirac, Proc. Roy. Soc. Lond. 109, 642 (1925).

³ E. Schrödinger, Ann. d. Phys. (4) 79, 361, 489, 734 (1926); 80, 437 (1926); 81, 109 (1926), Collected Papers on Wave Mechanics, Blackie and Son Ltd., London (1928).

* Schrödinger followed de Broglie's idea of matter waves in setting up his equation. Later he proved the equivalence of his approach to that of Heisenberg's. See ref. 4 above.

⁵ W. Heisenberg, Z. Physik 43, 172 (1927).

⁶ N. Bohr, Naturwiss. 16, 245 (1928) (Also printed as Appendix II in A.a.N.).

† i.e. phenomena concerning Nuclear, Atomic and Molecular Physics.

a path difference of an integral number of the wave-length between the light waves emerging from the various lines of the grating. We can assume, on the basis of the *superposition principle*, which is supported by a large set of experimental data, that this result of the wave theory corresponds to reality. Indeed, it holds also for arbitrarily weak intensities of the incident radiation (which is typical of this kind of phenomena) and hence for an individual atom emitting light. From the corpuscular standpoint, this process would be represented as follows: First an emission occurs from the atom; then (according to the relevant time of propagation of light) a scattering process connected with an observable recoil takes place at the diffraction grating; and finally an absorption process ensues at the point considered. The fact that light can reach only such points behind the grating lying along definite directions (which can be calculated using wave theory) of the diffracted photon requires the presence of *all* the atoms of the diffraction grating. The assumption that it would be possible to fix the point on the diffraction grating at which it is hit by the photon, without altering the nature of the diffraction pattern, will lead to insurmountable difficulties. The behaviour of the photon must be determined at every instant of time from the locations of all the existing atoms but in this case the theory of classical wave fields will not be adequate for predicting the statistical behaviour of the photon. As will be explained presently, there is no wave field with the following properties: (i) its intensity vanishes at all points of the grating with the exception of a single line and (ii) only specified directions of the diffracted rays can appear in it. It is only possible to realise either one or the other of these two properties through a wave field. Therefore, in order to avoid a conflict with the superposition principle, we must necessarily make the following assumption: If a definite line of the grating is hit by the photon (and none else) then there can be no influence of the other lines on the diffraction pattern behind the grating. Thus, this diffraction pattern must be the same as if only a single line of the grating is hit by the photon.

The above requirement is not, of course, tied up with the particular form of the diffraction experiment under consideration but can be generalised to all possible interference experiments. All such experiments, in general, have the common feature that light waves (from the same source), which have traversed different paths and hence have a phase difference, meet again at another point. We have to postulate that if a photon has taken, in any particular case, one of these paths, then there is no possibility of observing the interference pattern predicted by wave theory (cf. Sec. 16).

As already mentioned, this requirement is contained in another requirement which is more general and which can be formulated in the following way: *All the properties (possibly statistical) arising from other measurements made (earlier or later) on a photon, which follow from a knowledge of the result of a particular measurement, can be uniquely determined by statements on the wave field associated with this particular measurement.* We have to impose on this wave field the condition that it can always be generated by a superposition of plane waves in different directions and with different wave-lengths. We then speak of a wave-packet. Even without analysing the results of measurement made on a photon, we can express the fact that the photon finds itself in a certain space-time region in terms of the wave-packet associated with it. This we do by saying that the wave amplitudes are significantly

different from zero only within the concerned space-time region. We denote the (complex) phase of a plane wave by

$$e^{i(\sum_i k_i x_i - \omega t)}, \quad (1.1)$$

where the vector \vec{k} with components k_i , is in the direction of the wave normal and has the magnitude $2\pi/\lambda$, where λ is the wavelength. \vec{k} is the wave number vector; ω signifies the "angular frequency" or 2π times the frequency of the oscillation, v . The frequency ω is a function of k_1, k_2, k_3 , and is uniquely determined by the nature of the waves. In fact, for electromagnetic waves in vacuum, we have

$$\sum_i k_i^2 = \frac{\omega^2}{c^2}, \quad (1.2)$$

where c is the velocity of light in vacuum. It is, however, important to note that the following conclusions are independent of the special form of the function $\omega(k_1, k_2, k_3)$. For the most general wave field, each component of any arbitrary field strength can be represented by

$$u(\vec{x}, t) = \int A(\vec{k}) e^{i((\vec{k} \cdot \vec{x}) - \omega t)} d^3k, \quad (1.3)$$

where $A(\vec{k})$ denotes a function of k_1, k_2 , and k_3 . By elementary considerations, it can be shown that if $u(\vec{x}, t)$, for fixed t , is different from zero only inside a spatial region with dimensions $\Delta x_1, \Delta x_2, \Delta x_3$ and, at the same time, $A(\vec{k})$ is different from zero only inside the region of "k-space" with dimensions $\Delta k_1, \Delta k_2$, and Δk_3 , then the three products $\Delta x_i \Delta k_i$, where $i = 1, 2, 3$ cannot be arbitrarily small, but must be at least of the order 1:

$$\Delta x_i \Delta k_i \sim 1. \quad (1.4)$$

We shall speak about the quantitative refinement of this principle and its proof later. An analogous law holds for the spread Δt in time t , within which $u(\vec{x}, t)$ is different from zero for a fixed space-point (x_1, x_2, x_3) and the spread $\Delta \omega$ of frequency, which belongs to the region of the k-space mentioned, where $A(\vec{k})$ is essentially non-zero. Here again, we have the relation

$$\Delta \omega \Delta t \sim 1. \quad (1.4')$$

From the condition (1.4) it follows immediately that for a wave-packet of width equal to the distance between two lines of the grating, the angular width of the diffracted pencil of rays is large enough to cover (at least) two successive diffraction maxima and hence the diffraction pattern is effaced.

Since the measurements with a photon always involve its interaction with matter, we can draw conclusions about the material bodies from conditions (1.4) and (1.4'), which are essential for retaining the corpuscular picture in interference phenomena. The concept of a quantum of light is introduced in order to calculate the exchange of energy and momentum between light and matter, assuming that the laws of conservation of momentum and energy strictly hold for this exchange (and it is only through these conservation laws that energy and momentum are really defined). This exchange would be correctly described if one attributes to the photon a momentum \vec{p} in its direction of propagation with magnitude $\hbar \frac{\omega}{c}$ and an energy equal to $\hbar \omega$ where the fundamental constant \hbar denotes Plank's constant \hbar divided by 2π .*

* The notation, $\hbar = h/2\pi$ is due to Dirac.

Remembering the definition of the vector \vec{k} and eq. (1.2), this statement can be expressed by the relations

$$\vec{p} = \hbar \vec{k}; \quad E = \hbar \omega. \quad (\text{I})$$

The relations (1.4) and (1.4') lead to the results: (i) The position of the photon at a particular time cannot be determined simultaneously with its momentum and (ii) the energy of the photon cannot be determined simultaneously with the instant of time when it passed a particular spatial position. Indeed we have

$$\Delta p_i \Delta x_i \sim \hbar; \quad \Delta E \Delta t \sim \hbar. \quad (\text{II})$$

These are the uncertainty relations first established by Heisenberg; the derivation given here is due to Bohr. In a scattering process such as that between the photon and a material particle the interaction can take place as soon as the two particles meet together in space and time so that Δx and Δt are the same for both. Suppose we are able to measure p_i and E of the material particle both before and after the interaction with greater accuracy than what is implied by condition (II). Then, using the conservation laws, we can obtain more exact information regarding Δp_i and ΔE (for the photon also), than that corresponding to the condition (II). *Thus, if this relation is to hold for the photon and also if the conservation laws of energy and momentum are to be preserved for the interaction of the photon with material particles, then these uncertainty relations must be valid in general, not only for photons, but also for all material particles* (i.e. for electrons and protons as well as for macroscopic bodies).

The simplest interpretation of the above *general* limitation on the applicability of the classical particle picture (to which we are led in this way) consists in the assumption that even ordinary matter possesses wave-like properties, and that the wave number vector and the frequency of the corresponding wave are determined by the universal relation (I). *The existence of a duality between waves and particles and the validity of relation (I) for matter also*, forms the content of de Broglie's hypothesis of matter waves, which has received brilliant confirmation* through experiments on the diffraction of (charged and neutral) matter waves on crystal lattices.

The necessity for the universality of the wave-corpuscle duality for a correct description of phenomena can be illustrated with the example (discussed above) of the diffraction of a photon by a grating. To start with, let us imagine that the position at which the photon hits the grating can be determined approximately in the following way. We imagine certain parts of the grating to move with respect to one another and identify the part experiencing a recoil by the photon hitting it. Such an experimental arrangement is, in fact, possible but it is not correct to say that the diffraction pattern will be the same as when the parts of the grating are rigidly bound to one another. The momentum of the part of the grating in question must (before the photon strikes it) be subject to an uncertainty which is certainly smaller than the recoil momentum Δp_i , in order that the latter can be observed. Now the wave nature of the parts of the grating becomes important and, hence, by (II), there arises an uncertainty $\Delta x_i > (\hbar / \Delta p_i)$ in the relative positions of the movable parts of the

* C. Davisson and Germer, Nature, 119, 558 (1927); Phy. Rev. 30, 705 (1927); Kikuchi, Japan J. Phys., 5, 83 (1928); G.P. Thomson, Nature, 120, 802 (1927); Proc. Roy. Soc. London, A 117, 600 (1928).

grating. This uncertainty is precisely of such a magnitude that the resulting diffraction pattern will be the same as if only the part of the grating hit by the photon was present during the interaction.

All that has been said till now about the diffraction of photons holds also for the diffraction of matter waves. Only, the relation between frequency and wave number, which for photons was given by (1.2), is quite different for matter waves. According to the relativistic mechanics of a point-mass, we have the relation between energy and momentum

$$\frac{E^2}{c^2} = m^2 c^2 + \sum_i p_i^2, \quad (1.5)$$

where m denotes the rest-mass of the particle.

From (I) we have for the waves:

$$\frac{\omega^2}{c^2} = \frac{m^2 c^2}{\hbar^2} + \sum_i k_i^2 = \frac{\omega_0^2}{c^2} + \sum_i k_i^2, \quad (1.5')$$

where

$$\omega_0 = \frac{mc^2}{\hbar}. \quad (1.6)$$

The relations (I) between energy-momentum and frequency-wave number vector are *relativistically invariant*, since both $(\vec{p}, i(E/c))$ and $(\vec{k}, i(\omega/c))$ form the components of four vectors; similarly, the relations (1.5) and (1.5') are invariant. For $m = 0$ the eqs. (1.5) and (1.5') transform into the corresponding ones for the energy and momentum of a photon.

Not only the energy and momentum of a particle, but also *the velocity of the particle* can be connected with a simple quantity characteristic of the associated waves. This quantity (as de Broglie has shown) is equal to the *group velocity of the waves*. In fact, the velocity of the particle is determined by⁷

$$dE = \sum_i v_i d\dot{p}_i$$

or

$$v_i = \frac{\partial E}{\partial \dot{p}_i}. \quad (1.7)$$

and the group velocity of the waves by

$$v_i = \frac{\partial \omega}{\partial k_i}. \quad (1.7')$$

Both expressions are consistent with each other on account of relation (I). This situation is important in view of the fact that in cases where diffraction effects can be neglected, the wave-packets themselves move along the classical paths. In the case of the free particle considered here, the motion is rectilinear (cf. Sec. 4). From the relativistic relation (1.5) we have

$$v_i = \frac{\partial E}{\partial \dot{p}_i} = \frac{c^2 p_i}{E}. \quad (1.5a)$$

⁷ It is to be noted that the expression for the group velocity also yields the correct relationship between the phase velocity and ray velocity in the case of dispersive crystals. Since the wave normal and ray do not have the same directions in this case v is no longer parallel to \vec{k} , but the relation (1.7') is still valid.

Hence $p_i = (E/c^2)v_i$ and on substituting for p_i in (1.5) we get

$$\left. \begin{aligned} \frac{E^2}{c^2} \left(1 - \frac{v^2}{c^2} \right) &= m^2 c^2, \\ E &= \sqrt{m^2 c^2 + \sum_i p_i^2}, \\ p_i &= \frac{mv_i}{\sqrt{1 - v^2/c^2}}. \end{aligned} \right\} \quad (1.5b)$$

These are the well-known expressions for energy and momentum in terms of the velocity.

In the *non-relativistic case* $|p| \ll mc$, which is of great importance for what follows, we have

$$\frac{E}{c} = \sqrt{m^2 c^2 + \sum_i p_i^2} = mc \left(1 + \frac{1}{2m^2 c^2} \sum_i p_i^2 \right)^{1/2}$$

or

$$E = mc^2 + \frac{1}{2m} \sum_i p_i^2, \quad (1.8)$$

and

$$\omega = \omega_0 + \frac{\hbar}{2m} \sum_i k_i^2. \quad (1.8')$$

We note further that we have taken here the positive square roots of E and ω which is in agreement with experience. But

$$E = - \left(m c^2 + \frac{1}{2m} \sum_i p_i^2 \right) \quad (1.9)$$

is also possible (in principle). We shall come back to this in more detail in Section 18. If we, however, restrict ourselves to the first possibility, it is convenient to shift the zero point of the energy scale:

$$E' = E - mc^2; \quad \omega' = \omega - \omega_0. \quad (1.10)$$

Then we have

$$\left. \begin{aligned} E' &= \frac{1}{2m} \sum_i p_i^2, \\ \omega' &= \frac{\hbar}{2m} \sum_i k_i^2, \\ v_i &= \frac{p_i}{m} = \frac{\hbar k_i}{m}. \end{aligned} \right\} \quad (1.11)$$

Hence

$$\lambda = \frac{2\pi}{|\hbar|} = \frac{2\pi\hbar}{mv}, \quad (1.12)$$

where v denotes the magnitude of the velocity. This is the celebrated formula of de Broglie for the wave-length of matter waves.

The uncertainty relations (II) for material particles show that even for a free particle classical kinematics cannot be used without restriction. For, these relations contain the statement that an exact knowledge of the position of the particle results in the

complete impossibility of determining (and not merely the ignorance of) the momentum of the particle and vice versa. The distinction between *indeterminacy* and *ignorance* and the connection between the two concepts are decisive for the whole of quantum mechanics. This may be further explained through the example of an experimental arrangement, in which a photon can pass through two holes producing (in the sense of a statistical average by repetition of the experiment) a diffraction pattern on a screen placed behind the holes. In this case it is not definite through which hole the photon has passed. If, on the other hand, we have an experimental arrangement in which only one of the holes is open to the photon, but if we do not know which of the two holes is open, then we say: It is not known through which hole the photon has passed. Evidently, the diffraction figure in this case results from the addition of the intensities of the diffraction patterns for individual holes. These intensities are to be added after multiplication by suitable weight factors. Generalising, we can say: *Due to the indeterminacy in the property of a system prepared in a specific manner (i.e. in a definite state of the system), every experiment for measuring the property concerned destroys (at least partly) the influence of a prior knowledge of the system on the (possibly statistical) statements about the results of future measurements.* Hence, it is correct to say that here the measurement of the system leads to a new state. Moreover, a part of the effect, transferred from the measuring apparatus to the system, is itself left undetermined.

Hence, in order to determine the position and momentum of a particle simultaneously, *mutually exclusive experimental arrangements must be made use of.* For determining the position of the particle, spatially fixed apparatuses (scales, clocks, screens) are always necessary (to which an indeterminate amount of momentum is transferred); on the other hand, the determination of momentum would prevent the pinpointing of the particle in space and time. It would also not be of any use if we had determined this position before. The influence of the apparatus for measuring the momentum (position) of the system is such that within the limits given by the uncertainty relationships the possibility of using a knowledge of the earlier position (momentum) for the prediction of the results of later measurements of the position (momentum) is lost. If, due to this, the use of a classical concept excludes that of another, we call both concepts (e.g., position and momentum co-ordinates of a particle) *complementary* (to each other), following Bohr. We might call modern quantum theory as "The Theory of Complementarity" (in analogy with the terminology "Theory of Relativity").

We shall see that "complementarity" has no analogue in the classical theory of gases which also uses statistical laws.⁸ This theory of gases does not contain the assertion (made valid by the finiteness of the quantum of action) that by making measurements on a system the knowledge obtained by earlier measurements on the same must be lost under certain circumstances, i.e. that this knowledge can no longer be retrieved. (This statement also underlines the essential difference between quantum mechanics and the old quantum theory of Bohr, Kramers, and Slater.) As

⁸ However, N. Bohr refers in his Faraday lecture, J. Chem. Soc. (1932), 349, particularly pp. 376 and 377, to the fact that even in classical statistical mechanics, one can speak, in a somewhat different sense, of the complementarity of the knowledge of the microscopic molecular motion, on the one hand, and the macroscopic temperature of the system, on the other.

already mentioned, the possibility of viewing physical phenomena objectively and hence also the possibility of their causal space-time description is lost thereby. If these phenomena are to be properly described, then we have to make a *choice* (by observation) which is independent of the system being described, the point of separation between the observer and the observed being left arbitrary (see Sec. 9).

In the following it will be expounded as to how the *statistical* characterisation of the states and the appropriateness of using a statistical interpretation can be satisfactorily established.

2. The Measurement of Position and Momentum

For a closer characterisation of the state of a material particle it is first necessary to investigate the meaning that can be attached to the concepts of position and momentum of a particle even outside the region of validity of classical mechanics. As regards the position of the particle, we use, in order to fix it, an effect by which the particle can act only when the latter is found at a definite point. Luckily we have in the scattering of light such an effect which is displayed both by charged elementary particles and by macroscopic bodies. Let us imagine, for example, that the (x, y) plane is irradiated by a wave-train of limited length such that a particular point (x_0, y_0) of this plane will be irradiated at time t_0 . Time t_0 has a spread Δt , which cannot be less than $1/v$, where v is the mean frequency of the light. By using light of the shortest possible wave-length, Δt can be made as small as possible. We can further imagine the intensity of the light to be so great that at least one photon will be certainly scattered by the particle if the pencil of light passes by it. One can now use any optical magnifying instrument (camera obscura, telescope, microscope) to make a more refined determination of the position of a material particle by a rough macroscopic measurement of the point at which the effect of a scattered quantum is felt. For this purpose it suffices to observe an individual photon. The limits on the accuracy of the measurement of position are always determined by the boundaries of the optical images, in so far as these images are restricted by the diffraction effects given by classical wave optics. It is well known, for instance, for a microscope that the limits on the accuracy of the location of the image is given by

$$\Delta x \sim \frac{\lambda'}{\sin \epsilon}, \quad (2.1)$$

where λ' denotes the wave-length of the scattered radiation (λ' can be different from the wave-length of the incident radiation) and ϵ denotes the half angle of aperture of the objective. The direction of the scattered light must then in principle be considered as undetermined within this angle ϵ ; hence the component of the momentum of the material particle in the x -direction is undetermined, after the collision, by an amount

$$\Delta p_x \sim \frac{\lambda}{\lambda'} \sin \epsilon. \quad (2.2)$$

From eqs. (2.1) and (2.2), the uncertainty relation

$$\Delta p_x \Delta x \sim \hbar$$

can easily be checked. We shall, however, discuss further the accuracy with which it is possible to determine the position of a particle in the *gedanken* (thought) experiment discussed above. According to eq. (2.1) it is evidently advantageous to make the wave-length of the scattered light as small as possible. If the wave-length of the scattered radiation is equal to that of the incident radiation, the accuracy of the measurement of position could be arbitrarily enhanced since the wave-length could be made arbitrarily small. At the same time, as already mentioned above, the instant of time at which the position is measured can be made to lie in an arbitrarily small interval. On account of Compton effect there is a change in the frequency of the scattered radiation which is determined by the energy and momentum laws. This has the consequence that even in the limit $v \rightarrow \infty$ ($\lambda = (c/v) \rightarrow 0$) the frequency v' of the scattered radiation cannot exceed a finite value. If \vec{p} and $E = c \sqrt{m^2 c^2 + \vec{p}^2}$ are the momentum and energy of a material particle before the scattering process, then in this limit ($v \rightarrow \infty$) corresponding to a maximum for v' and hence a minimum for $\lambda' = c/v'$, we have

$$\left. \begin{aligned} v' &\sim \frac{E}{h} = \frac{m_0 c^2}{h} \frac{1}{\sqrt{1 - v^2/c^2}}, \\ \lambda' &\sim \frac{\hbar c}{E} = \frac{\hbar}{mc} \sqrt{1 - \frac{v^2}{c^2}}. \end{aligned} \right\} \quad (2.3)$$

(Here very small scattering angles are excluded because they are not suited for measuring position¹ on geometrical grounds.) For maximum accuracy in determining the position of a particle *by means of the experiment of the scattering of a photon by an optical instrument (discussed here)* it follows that

$$\left. \begin{aligned} \Delta x &\sim \frac{\hbar c}{E} = \frac{\hbar}{mc} \sqrt{1 - \frac{v^2}{c^2}}, \\ \Delta t &\sim \frac{1}{v'} \sim \frac{\hbar}{E} = \frac{\hbar}{mc^2} \sqrt{1 - \frac{v^2}{c^2}}. \end{aligned} \right\} \quad (2.4)$$

The second line in eq. (2.4) follows from the fact that the duration of the scattering process, i.e. the time, within which an interaction between the photon and the material particle takes place, can never be significantly smaller than the periods of oscillation of the incident and scattered radiation. This duration of the time of measurement of position is important because it determines also whether these results of measurement can be used for predictions of later measurements of position. The possibility of the measurement of position at a later instant of time arises in the following sense. If one redetermines the position after a lapse of time τ , the result of this determination is indeed not predictable in individual cases, however, as an average over several repeated experiments, we will indeed find a certain mean position $\bar{x}(t_0 + \tau)$ with a given mean error, $\Delta = \sqrt{(\Delta x)^2}$. Then $\bar{x}(t_0 + \tau) - \bar{x}(t_0)$ and $\Delta(t_0 + \tau) - \Delta(t_0)$ can be made arbitrarily small by making τ small. If the instant of

¹ If the incident radiation is directed opposite to the original direction of motion of the particle, while the scattered light is parallel to it, it follows, e.g., from the energy and momentum conservation laws that

$$v' = v \frac{E + c p_x}{2hv - c p_x + E};$$

hence for $h v \gg E$

$$v' \sim \frac{E + c p_x}{2h} = \frac{1}{2} \left(1 + \frac{v_x}{c}\right) \frac{mc^2}{h} \frac{1}{\sqrt{1 - v^2/c^2}}$$

time for the first determination of position had remained completely undetermined, its value could not be used for predicting the results of another measurement of position, and would have been, in this sense, of no physical consequence.

The limit of accuracy given by (2.4) for the determination of position is of importance, at the most, for atomic nuclei and electrons, since for an atom as a whole the dimension, in general, is larger than the Compton wave-length h/mc . Whether this limit for the first-named particles (nuclei, electrons) has a significance² or whether it can be reached by indirect methods cannot be decided beforehand by elementary considerations. It is completely dependent on the foundation on which a relativistic quantum mechanics can be successfully built up. We do not consider this point here in order not to complicate the problem too much and to avoid transgressing the domain of our present knowledge. In particular, the atomic constitution of the scales and clocks themselves has not been taken into account. Hence, the possible constraints due to the presence of arbitrarily small screens, lenses, or mirrors are deliberately ignored. We wish to stress here that the concept of the position of a material particle at a particular time has a meaning even outside the region of validity of classical mechanics. The measurement of position with great accuracy is always possible since the wave-length of the material particle amounts to

$$\lambda_m = \frac{\hbar}{|p|} = \frac{\hbar}{m v} \sqrt{1 - \frac{v^2}{c^2}}$$

and according to (2.4)

$$\Delta x \sim \lambda_m \cdot \frac{v}{c}. \quad (2.5)$$

At least in non-relativistic quantum mechanics, where $v \ll c$, the following basic assumption is, therefore, natural: *In every state of a system and indeed for free particles there exists, at each instant of time t , a probability $W(x_1, x_2, x_3; t)$ $dx_1 dx_2 dx_3$ that the particle is found in the volume interval $(x_1, x_1 + dx_1; x_2, x_2 + dx_2; x_3, x_3 + dx_3)$.*

This basic assumption is neither self-evident nor is it a direct consequence of the uncertainty relations (II). This is clear, since for a photon such a statement regarding its position is not meaningful outside the purview of classical geometrical optics (this will be discussed later, see Sec. 25d). The position of the photon cannot be determined more accurately than the wave-length of light and the time required for this determination cannot be shorter than its period of oscillation. Hence, there is no photon-density with properties similar to that of the density of material particles.³ As will be made clear in the following, the analogy between light and matter cannot be pushed as far as it originally appeared possible. The analogy is confined, on the other hand, completely to the fundamental relations (I) between energy-momentum and frequency-wave-length, which are valid for both photons and material particles.

In the formulation of the basic assumption above, a distinction is made between space and time since the position co-ordinates are considered within a margin dx , while the time co-ordinate is exactly fixed.⁴ Actually, as we have seen, this time-point

² This standpoint was taken by L. Landau and R.F. Peierls, Z. Physik **69**, 56 (1931).

³ In the literature and even in some textbooks many incorrect statements are made on this point.

⁴ For this standpoint, one can in particular refer to E. Schrödinger (Berl. Ber. (1931) p. 238). In this connection, it is stressed there that an ideal clock, i.e. one which gives the time exactly, will possess an infinitely large uncertainty in energy and hence also an infinite energy. According to us, this does not

cannot be fixed more accurately than $\Delta t = \Delta x/c$, if the order of magnitude of the error in position-determination is Δx . Only in the limiting case of non-relativistic quantum mechanics, in which c can be treated as infinitely large, so to say, does it appear a meaningful idealisation to neglect the time interval Δt for fixed Δx , and set Δt mathematically equal to zero.

We now proceed to discuss the problem of determining the momentum of particles. According to Bohr, the scattering of a photon by a material particle can be utilised, since the Doppler effect in a particular direction of the scattered radiation (along with the frequency and direction of the incident radiation) allows us to arrive at a conclusion about the velocity of the material particle. Since the accuracy of the determination of v' is limited by the finite duration T of the interaction between light and matter according to

$$\Delta v' = \frac{1}{T} \quad (2.6)$$

it is advantageous in this case – in contrast to the case of the determination of position – to choose the duration of time to be large. If we consider for the sake of simplicity the case in which the material particle is moving in the $+x$ direction before the process, so that $p_x = p_z = 0$ and in which the radiation incident on the particle in the $-x$ direction is scattered along the $+x$ direction, then we have

$$-\frac{\hbar v}{c} + p_x = p'_x + \frac{\hbar v'}{c}$$

or

$$p'_x = p_x - \frac{\hbar v}{c} - \frac{\hbar v'}{c} \quad (2.7)$$

and

$$\hbar v - \hbar v' = E' - E. \quad (2.8)$$

Since v is given, an exact knowledge of p_x (and p'_x) would follow from an exact knowledge of v' . To find the connection between the inaccuracy Δp_x of p_x and the inaccuracy $\Delta v'$ of v' , we have first to calculate $\partial v'/\partial p_x$ from (2.8). [According to (2.7) p'_x is to be thought of as a function of p_x and v' , whereas v is fixed.] Remembering

$$\frac{\partial E'}{\partial p_x} = v'_x; \quad \frac{\partial E}{\partial p_x} = v_x$$

(which relation is also valid in the relativistic case), we find

$$-\hbar \frac{\partial v'}{\partial p_x} = v'_x \left(1 - \frac{\hbar}{c} \frac{\partial v'}{\partial p_x}\right) - v_x; \quad \hbar \frac{\partial v'}{\partial p_x} \left(1 - \frac{v'_x}{c}\right) = v_x - v'_x.$$

The errors are related as follows:

$$\hbar \Delta v' = \frac{(v_x - v'_x)}{1 - v'_x/c} \Delta p_x.$$

Now for small v , v'_x is nearly equal to v_x ; for increasing v , v'_x decreases and will be finally negative becoming $-c$ for very large v . The denominator $1 - (v'_x/c)$, therefore,

mean that the use of the usual concept of time contradicts quantum mechanics, since such an ideal clock can be approximated arbitrarily. One can imagine, e.g., a very short (in the limit, infinitesimally short) wave train of light, which due to the presence of a suitable mirror describes a closed path. But the question of the existence of such a mirror is outside our purview, as already mentioned in the text.

changes from $1 - (v_x/c)$ to 2 . Considering only the order of magnitude, we have

$$\hbar \Delta p' \sim (v_x - v'_x) \Delta p_x. \quad (2.9)$$

Hence, according to eq. (2.6)

$$\Delta p_x \sim \frac{\hbar}{(v_x - v'_x) T}. \quad (2.10)$$

On the other hand, for the uncertainty in the position of the particle after the process

$$\Delta x \sim (v_x - v'_x) T,$$

since the time-point within the interval T when the particle changes its velocity remains undetermined. In this way the uncertainty relation

$$\Delta p_x \Delta x \sim \hbar$$

is again verified. Equation (2.10) shows that the momentum of the particle can be determined even in an arbitrarily short time, if the (known) change in the velocity of the particle during the process could be arbitrarily large. Actually, it cannot be greater than $2c$, so that considering orders of magnitude, we have

$$\Delta p_x \sim \frac{\hbar}{c \Delta t}. \quad (2.11)$$

Here Δt is written in place of T to denote that T gives at the same time the uncertainty in the time-point at which the momentum value p_x was realised. Moreover, the results, (2.10) and (2.11), considered as the lower limits of the error Δp_x are independent of any special assumptions about the direction of radiation and the velocity of the material particle.

In the case of free particles, the restriction on the accuracy of the determination of momentum due to finite duration of time T is not essential, since the momentum of the particle is now constant in time. We can, therefore, assume that *for every state of a system and indeed for a free particle, there exists a probability $W(p_1, p_2, p_3)$ $dp_1 dp_2 dp_3$ that the momentum of the particle lies in the volume interval $(p_1, p_1 + dp_1; p_2, p_2 + dp_2; p_3, p_3 + dp_3)$.* (In the case of free radiation also, this assumption evidently proves to be correct for a photon.*)

In addition to the limitation in accuracy given by (2.10), the measurements of momentum are not in general repeatable since a possibly large but known change in momentum can take place. Only when the duration of time T of measurement is chosen so long that for given Δp_x , $(p'_x - p_x)$ can also be made small (incident light of long wave-length), will a second measurement of momentum made immediately after the first lead to the same result. However, in all cases (even for measurements of short duration) the result of the second measurement of momentum can be predicted on the basis of the first. These facts are important for discussing the question of measurement of the momentum of bound particles, since, as we shall see, for these particles, only a limited time is available for the measurement.

*See, e.g., A.I. Akhiezer and V.B. Berestetskii, Quantum Electrodynamics, Interscience Publishers, New York (1965).

CHAPTER II

Schrödinger Equation and Operator Calculus

3. The Wave Function of Free Particles

We now introduce the basic assumptions regarding the position probabilities $W(x_1, x_2, x_3)$ and the momentum probabilities $W(p_1, p_2, p_3)$ of a particle, which are in agreement with the uncertainty relations (II) and the wave nature of matter. We shall restrict ourselves to the non-relativistic region, where the velocity of the particle is small compared to that of light and the frequency of the waves is related to the phase vector \vec{k} by eq. (1.8'):

$$\omega = \omega_0 + \frac{\hbar}{2m} \sum_i k_i^2. \quad (1.8')$$

By restricting ourselves to the non-relativistic case, we are excluding the photons from consideration. We shall consider the relativistic theory in Chapters IX and X. We imagine, therefore, the functions

$$\psi(\vec{x}, t) = \frac{1}{\sqrt{(2\pi)^3}} \int A(\vec{k}) e^{i(\vec{k}\cdot\vec{x}) - \omega t} d^3 k \quad (3.1)$$

to be constructed as in (1.3) where $|\vec{k}|$ and ω satisfy (1.8') and hence are positive. The reason for introducing the factor $1/\sqrt{(2\pi)^3}$ will be given later. Further we construct the complex conjugate function

$$\psi^*(\vec{x}, t) = \frac{1}{\sqrt{(2\pi)^3}} \int A^*(\vec{k}) e^{-i(\vec{k}\cdot\vec{x}) - \omega t} d^3 k. \quad (3.1'')$$

Now introducing, instead of \vec{k} and ω , the momentum $\vec{p} = \hbar\vec{k}$ and the energy $E = \hbar\omega$ of the particle according to I in (3.1) and (3.1''), these functions can also be written as

$$\psi(\vec{x}, t) = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int A(\vec{p}) e^{\frac{i}{\hbar}[(\vec{p}\cdot\vec{x}) - Et]} d^3 p, \quad (3.1')$$

$$\psi^*(\vec{x}, t) = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int A^*(\vec{p}) e^{-\frac{i}{\hbar}[(\vec{p}\cdot\vec{x}) - Et]} d^3 p. \quad (3.1'')$$

The functions $A(\vec{p})$ and $A(\vec{k})$ differ by a numerical factor such that

$|A(\vec{p})|^2 dp_1 dp_2 dp_3 = |A(\vec{k})|^2 dk_1 dk_2 dk_3$. Setting

$$\varphi(\vec{p}) = A(\vec{p}) e^{-\frac{i}{\hbar} Et}, \quad (3.2)$$

$\varphi(\vec{p})$ satisfies the equation

$$-\frac{\hbar}{i} \frac{\partial \varphi}{\partial t} = E \varphi = \left(E_0 + \sum_i \frac{p_i^2}{2m} \right) \varphi. \quad (3.3)$$

Then ψ and ψ^* can be written as

$$\psi(\vec{x}, t) = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int \varphi(\vec{p}) e^{\frac{i}{\hbar} (\vec{p} \cdot \vec{x})} d^3 p, \quad (3.1'')$$

$$\psi^*(\vec{x}, t) = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int \varphi^*(\vec{p}) e^{-\frac{i}{\hbar} (\vec{p} \cdot \vec{x})} d^3 p. \quad (3.1''')$$

According to the Fourier integral theorem, the inverse forms of these relations are

$$\varphi(\vec{p}) = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int \psi(\vec{x}, t) e^{-\frac{i}{\hbar} (\vec{p} \cdot \vec{x})} d^3 x \quad (3.4)$$

$$A(\vec{p}) = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int \psi(\vec{x}, t) e^{-\frac{i}{\hbar} [(\vec{p} \cdot \vec{x}) - Et]} d^3 x \quad (3.4')$$

$$A(\vec{k}) = \frac{1}{\sqrt{(2\pi)^3}} \int \psi(\vec{x}, t) e^{-i[(\vec{k} \cdot \vec{x}) - Et]} d^3 x. \quad (3.4'')$$

Further we have the completeness relation

$$\int \psi^* \psi d^3 x = \int \varphi^* \varphi d^3 p = \int A^* A d^3 p, \quad (3.5)$$

which also represents the basis for the choice of the numerical factors in (3.1) and (3.1').

It is easy to see, according to (1.8'), that these functions satisfy the differential equations

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \left(E_0 - \frac{\hbar^2}{2m} \Delta \right) \psi, \quad (3.6)$$

$$+\frac{\hbar}{i} \frac{\partial \psi^*}{\partial t} = \left(E_0 - \frac{\hbar^2}{2m} \Delta \right) \psi^*, \quad (3.6'')$$

where, as in eq. (1.6), we have set

$$E_0 = \hbar \omega_0 = mc^3$$

and where Δ denotes the Laplacian operator. Conversely, (3.1) is the general¹ solution of the differential equation (3.6), if for every wave appearing in eq. (3.1) the relation (1.8') is satisfied. This relation arises, according to (I), from

$$E = E_0 + \frac{1}{2m} \sum_i p_i^2 \quad (1.8)$$

¹ If this expression is to cover the case involving a sum over various plane waves besides the integral, certain singularities must be allowed in $A(\vec{k})$ and the integrals must then be understood in the Stieltjes' sense.

which connects the energy and momentum of a particle according to the classical particle mechanics. Formally (3.6) results directly from (1.8), if we introduce the operators²

$$\mathbf{E} = -\frac{\hbar}{i} \frac{\partial}{\partial t}; \quad \mathbf{p}_i = \frac{\hbar}{i} \frac{\partial}{\partial x_i} \quad (3.7)$$

acting on the space-time functions and then replace (1.8) by the operator equation

$$\mathbf{E}\psi = \left(E_0 + \frac{1}{2m} \sum_i \mathbf{p}_i^2 \right) \psi \quad (3.8)$$

which is identical with (3.6). Equation (3.8) is similar to eq. (3.3).

In the following we shall have to deal with more general operators also, but all of them possess the property of *linearity*. By linearity we understand that D , the operator concerned, satisfies the condition

$$D(c_1\psi_1 + c_2\psi_2) = c_1 D\psi_1 + c_2 D\psi_2, \quad (3.9)$$

where c_1 and c_2 are two arbitrary constants and ψ_1 and ψ_2 denote arbitrary functions of certain variables. These variables can, under certain circumstances, assume, instead of continuous values, like the space-time co-ordinates, only discrete (or countably infinite sets of) values or even a finite number of values. The functions $D\psi$, however, are always understood to depend on the same variables as the functions ψ do. Returning to the operators (3.7), we observe that the association of precisely these operators to the energies and momenta leads to the Fourier integral relation (3.1'') between the space-time functions $\psi(\vec{x}, t)$ and the momentum functions, $\phi(\vec{p})$.

The functions introduced here acquire physical significance only when they are related to the probabilities, $W(\vec{k})$ and $W(\vec{p})$ for the energy and momentum of the particle in the concerned state. In this connection, it is essential to remember that these probabilities can *never be negative* and further we must have at every time-point

$$\int W(\vec{x}) d^3x = 1 \quad (3.10)$$

and

$$\int W(\vec{p}) d^3p = 1. \quad (3.10')$$

The simplest ansatz for $W(\vec{x})$, which satisfies these requirements is that $W(\vec{x})$ is a *definite quadratic form* constructed from the values of the functions ψ_p , ψ_p^* , \dots ($p = 1, 2, \dots$) each of which satisfies the eqs. (3.6) and (3.6*):

$$W(\vec{x}) = Q(\psi_p, \psi_p^*). \quad (3.11)$$

(Only the results will show that one can do without quartic forms or forms of higher degree.) In order to ensure that

$$\int W(\vec{x}) d^3x$$

is constant in time when (3.6) and (3.6*) are taken into account, we must necessarily set

$$Q(\psi_p, \psi_p^*) = \sum_p C_p \psi_p^* \psi_p \quad (3.12)$$

where C_p are real and positive numbers. We infer this form for Q both from (3.1) by

² Bold face letters denote, in general, operators in this book. Three-dimensional vectors are, therefore, denoted by an arrow over the letters.

means of the Fourier theorem and also from (3.6) by partial integration. For example, using integration by parts, we obtain immediately

$$\begin{aligned} -\frac{1}{2} \frac{\hbar}{i} \frac{\partial}{\partial t} (\psi^2) &= E_0 \psi^2 - \frac{\hbar^2}{2m} \sum_i \frac{\partial}{\partial x_i} \left(\psi \frac{\partial \psi}{\partial x_i} \right) + \frac{\hbar}{2m} (\text{grad } \psi)^2, \\ +\frac{1}{2} \frac{\hbar}{i} \frac{\partial}{\partial t} (\psi^{*2}) &= E_0 \psi^{*2} - \frac{\hbar^2}{2m} \sum_i \frac{\partial}{\partial x_i} \left(\psi^* \frac{\partial \psi^*}{\partial x_i} \right) + \frac{\hbar}{2m} (\text{grad } \psi^*)^2 \end{aligned}$$

and

$$\frac{\hbar}{i} \frac{\partial}{\partial t} (\psi \psi^*) = \frac{\hbar^2}{2m} \sum_i \frac{\partial}{\partial x_i} \left(\psi^* \frac{\partial \psi}{\partial x_i} - \psi \frac{\partial \psi^*}{\partial x_i} \right).$$

Hence, neither $\int \psi^2 d^3x$ nor $\int \psi^{*2} d^3x$ nor even any linear combination of them is constant in time, but still we have

$$\int \psi \psi^* d^3x = \text{const. (in time)} \quad (3.13)$$

(under the assumption that the surface integral arising from the partial integration over a very large sphere vanishes in the limit of an infinitely large region of integration). We note for future applications that the last of the differential equations written above assumes the form of a continuity equation

$$\frac{\partial \rho}{\partial t} + \text{div} \vec{i} = 0, \quad (3.14)$$

where we set $\rho = \psi^* \psi$ and

$$\vec{i} = \frac{\hbar}{2mi} (\psi^* \text{grad } \psi - \psi \text{grad } \psi^*). \quad (3.15)$$

If we stipulate that the time derivative of a real function is a new (different) function, we can say: Eq. (3.8) states that *a single real function is not sufficient in order to construct from wave functions of the form (3.1) a non-negative probability function that is constant in time when integrated over the whole space*.³ At least two real functions or one complex function and its conjugate are necessary for this purpose. The constants C_α can evidently be absorbed in ψ so that

$$W(\vec{x}) = \sum_\alpha \psi_\alpha^* \psi_\alpha = \sum_\alpha |\psi_\alpha|^2 \quad (3.12')$$

is the most general form for the probability $W(\vec{x})$. We shall see later that multi-component ψ -functions may be necessary if one is dealing with particles with spin. For the sake of simplicity, however, we shall disregard this for the time being and

³ This is connected with the fact that the real part $u = \frac{1}{2}(\psi + \psi^*)$ of ψ does not, according to (3.6) and (3.6*) satisfy a differential equation which is of the first order with respect to the time derivative, but satisfies only the "iterated" differential equation of the second order:

$$\left(-\frac{\hbar}{i} \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \Delta - E_0 \right) \left(+\frac{\hbar}{i} \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \Delta - E_0 \right) u = 0$$

or

$$\left[\hbar^2 \frac{\partial^2}{\partial t^2} + \left(\frac{\hbar^2}{2m} \Delta - E_0 \right)^2 \right] u = 0.$$

[For the imaginary part $v = (1/2i)(\psi - \psi^*)$ a similar statement holds.] A quadratic expression built from u , the volume integral of which is constant in time, must contain not only u and its space derivatives, but also the first derivative with respect to time.

work only with a single complex ψ -function so that

$$W(\vec{x}) = |\psi|^2 = \psi^* \psi \quad (3.12'')$$

with the normalisation condition

$$\int \psi^* \psi d^3x = 1. \quad (3.13')$$

According to the continuity equation (3.14) we can interpret the expression (3.15) as the *statistical current density* or the *probability current*. The vector $\vec{i}(x)$ is the probability that the particle passes through a unit surface element perpendicular to the x -axis, per unit time in the positive rather than in the negative x -direction.

Now it is easy to describe the probability density $W(\vec{p})$ in momentum space which (and not merely its integral) will be constant in time for free particles, since the momentum of the particle itself is constant. This probability density is given by

$$W(\vec{p}) = |A(\vec{p})|^2 = A^* A = \varphi^* \varphi. \quad (3.16)$$

One could perhaps think of defining $W(\vec{p})$ as

$$W(\vec{p}) = C(\vec{p}) |A(\vec{p})|^2$$

where, in general, $C(\vec{p})$ would be a positive function yet to be determined. On account of the completeness relation (3.5) following from (3.1'), it is, however, necessary to set $C(\vec{p}) \equiv 1$ since, from

$$\int W(\vec{x}) d^3x = 1$$

it must necessarily follow that

$$\int W(\vec{p}) d^3p = 1.$$

The statistical description of any state of a free material particle is thus completely given. Every such state is described by a wave-packet $\psi(\vec{x}, t)$ of the form (3.1) from which, according to eq. (3.4), the "wave packet" $\varphi(\vec{p})$ in momentum space follows uniquely. These functions $\psi(\vec{x}, t)$ and $\varphi(\vec{p})$, which are usually called "probability amplitudes" are *not*, however, *directly observable* with regard to their phases; this holds only for the probability densities $W(\vec{x})$ and $W(\vec{p})$. The complex wave function itself has only a *symbolic character* and serves to establish the connection between $W(\vec{x})$ and $W(\vec{p})$.⁴

From the foundations so far developed, we can draw several simple conclusions, which can be directly compared with experiment. In particular, we can construct the average values of any function of x or p and investigate their relationship as well as their temporal variation. For example

$$\bar{x}_i = \int x_i \psi^* \psi d^3x; \quad \bar{p}_i = \int p_i \varphi^* \varphi d^3p. \quad (3.17)$$

We shall also be interested in the mean extension of the wave-packet in ordinary space as well as in momentum space, which will be given by the "mean square deviations"

$$\overline{(\Delta x_i)^2} = \int (x_i - \bar{x}_i)^2 \psi^* \psi d^3x; \quad \overline{(\Delta p_i)^2} = \int (p_i - \bar{p}_i)^2 \varphi^* \varphi d^3p. \quad (3.18)$$

⁴ The mathematical problem, as to whether for given functions $W(\vec{x})$ and $W(\vec{p})$, the wave function ψ , if such a function exists is always uniquely determined (i.e. if $W(\vec{x})$ and $W(\vec{p})$ are physically compatible) has still not been investigated in all its generality.

The behaviour of the centre-of-mass of a wave-packet is obtained by using (3.1'') and (3.4) and by partial integration:

$$\begin{aligned}\bar{x}_t &= \frac{1}{\sqrt{(2\pi)^3}} \int x_i \psi^* d^3x \int \varphi(\vec{k}) e^{i(\vec{k} \cdot \vec{x})} d^3k = \frac{1}{\sqrt{(2\pi)^3}} \int \psi^* d^3x \int \varphi(\vec{k}) \frac{1}{i} \frac{\partial}{\partial k_i} (e^{i(\vec{k} \cdot \vec{x})}) d^3k \\ &= \frac{1}{\sqrt{(2\pi)^3}} \int \psi^* d^3x \int i \hbar \frac{\partial}{\partial p_i} [\varphi(\vec{p})] e^{i(\vec{p} \cdot \vec{x})} d^3p \\ &= \frac{1}{\sqrt{(2\pi\hbar)^3}} \int i \hbar \frac{\partial}{\partial p_i} [\varphi(\vec{p})] d^3p \int \psi^* e^{i(\vec{p} \cdot \vec{x})} d^3x = \int \varphi^*(\vec{p}) i \hbar \frac{\partial}{\partial p_i} [\varphi(\vec{p})] d^3p.\end{aligned}$$

Hence

$$\bar{x}_t = \int \varphi^*(\vec{p}) \cdot i \hbar \frac{\partial \varphi(\vec{p})}{\partial p_i} d^3p = \int \varphi^*(\vec{k}) \left(i \frac{\partial}{\partial k_i} \right) \varphi(\vec{k}) d^3k \quad (3.19)$$

or

$$\bar{x}_t = \int A^*(\vec{k}) e^{i\omega t} \left(i \frac{\partial}{\partial k_i} \right) [A(\vec{k}) e^{-i\omega t}] d^3k.$$

Finally we have

$$\bar{x}_t = \int A^* i \frac{\partial A}{\partial k_i} d^3k + t \int \frac{\partial \omega}{\partial k_i} A^* A d^3k. \quad (3.20)$$

Differentiating (3.20) with respect to time, we get

$$\frac{d\bar{x}_t}{dt} = \overline{\left(\frac{\partial \omega}{\partial k_i} \right)} = \overline{\left(\frac{\partial E}{\partial p_i} \right)} = \bar{v}_t = \frac{\bar{p}_i}{m}, \quad (3.21)$$

which represents the theorem on the group velocity. On the other hand, from the continuity eq. (3.14), it follows easily, on multiplication by x_i and partial integration, that

$$\frac{d\bar{x}_t}{dt} = \int i_t d^3x = \frac{1}{m} \int \psi^* \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial x_i} \right) d^3x. \quad (3.22)$$

Hence it follows by comparison with (3.21) that

$$\bar{p}_i = \int \varphi^*(\vec{p}) p_i \varphi(\vec{p}) d^3p = \int \psi^* \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial x_i} \right) d^3x, \quad (3.23)$$

which can be easily verified directly. The relations (3.19) and (3.23) can be further generalised. Let $F(x_i)$ be any rational integral function of x_i and $F(p_i)$ any rational function of p_i ; then

$$\overline{F(x_i)} = \int \psi^* F(x_i) \psi d^3x = \int \varphi^* F \left(i \hbar \frac{\partial}{\partial p_i} \right) \varphi d^3p, \quad (3.24)$$

$$\overline{F(p_i)} = \int \psi^* F \left(\frac{\hbar}{i} \frac{\partial}{\partial x_i} \right) \psi d^3x = \int \varphi^* F(p_i) \varphi d^3p. \quad (3.24')$$

These relations can be directly verified by partial integration and use of the Fourier integral theorem.⁵ Hence, e.g.,

$$\overline{p_i^2} = \int \varphi^* p_i^2 \varphi d^3p = \int \psi^* \left(-\hbar^2 \frac{\partial^2 \psi}{\partial x_i^2} \right) d^3x = +\hbar^2 \int \frac{\partial \psi^*}{\partial x_i} \cdot \frac{\partial \psi}{\partial x_i} d^3x,$$

$$\overline{x_i^2} = \int \psi^* x_i^2 \psi d^3x = \int \varphi^* \left(-\hbar^2 \frac{\partial^2 \varphi}{\partial p_i^2} \right) d^3p = +\hbar^2 \int \frac{\partial \varphi^*}{\partial p_i} \cdot \frac{\partial \varphi}{\partial p_i} d^3p.$$

⁵ For generalisations of this relation to functions other than rational integral functions, see Chapter IX, Sec. 18c and Chapter X, Sec. 25b.

To set up corresponding relations for $(\Delta x_i)^2$ and $(\Delta p_j)^2$, it is necessary to modify these equations only slightly. These are obtained most easily by making a transformation to a new frame of reference:

$$x' = x - x_0 - vt; \quad t' = t.$$

Of course, we use here the Galilean transformation, as we have provisionally neglected relativistic corrections. We also have

$$p'_x = p_x - mv; \quad E' = E - p_x v + \frac{m}{2} v^2.$$

In order to obtain the equation

$$-\frac{\hbar}{i} \frac{\partial \psi'}{\partial t} = E' \psi'$$

using (3.3) we set

$$\psi' = \psi e^{-\frac{i}{\hbar} [\frac{m}{2} v^2 - p_x v] t} e^{\frac{i}{\hbar} f},$$

where f , which is independent of t , is to be determined such that the function

$$\psi' = \frac{1}{\sqrt{(2\pi\hbar)^3}} \int \psi'(\vec{p}') e^{\frac{i}{\hbar} (\vec{p}' \cdot \vec{x}')} d^3 p'$$

has the property

$$W'(\vec{x}') = W(\vec{x})$$

or

$$\psi'^*(\vec{x}') \psi'(\vec{x}') = \psi^*(\vec{x}) \psi(\vec{x}).$$

This is achieved by setting $f = p_x x_0$. Then we obtain

$$\psi'(\vec{p}') = \psi(\vec{p}) e^{-\frac{i}{\hbar} [\frac{m}{2} v^2 t - p_x(x_0 + vt)]}, \quad (3.25)$$

$$\psi'(x') = \psi(x) e^{-\frac{i}{\hbar} [mv(x - x_0) - \frac{m}{2} v^2 t]} \quad (3.26)$$

or

$$\psi'(x', t') = \psi(x' + x_0 + vt') e^{-\frac{i}{\hbar} [mvx' + \frac{m}{2} v^2 t']}. \quad (3.26')$$

It is easily verified that this function satisfies the equation

$$-\frac{\hbar}{i} \frac{\partial \psi'}{\partial t'} = E_0 - \frac{\hbar^2}{2m} \Delta' \psi'.$$

Further we have for the current density (3.15)

$$\vec{j}' = \vec{i} - \vec{v} \psi^* \psi, \quad (3.27)$$

which obviously has an intuitive meaning.⁶

Since the mid-point of a wave-packet moves with constant velocity, according to (3.20), we can introduce a reference system K' , which itself moves along with the

⁶ V. Bargmann, Ann. of Math. 59, 1 (1954); particularly Sec. 6g gives a group-theoretic application of this.

mid-point of the packet so that the latter finds itself at rest at the origin of the new co-ordinate system. We have, then, in this new system

$$\bar{x} = 0; \quad \bar{p} = 0$$

and

$$\bar{x}^2 = \overline{(x - \bar{x})^2} = \overline{(\Delta x)^2}; \quad \bar{p}^2 = \overline{(p - \bar{p})^2} = \overline{(\Delta p)^2}.$$

The complete description would be $\bar{x}_i = 0, \dots$; in order to simplify the notation, we shall consider in the following the one-dimensional case and drop the prime. The mean value

$$\bar{p}^2 = \int p^* \varphi^* \varphi dp = \hbar^2 \int \frac{\partial \varphi^*}{\partial x} \cdot \frac{\partial \varphi}{\partial x} dx$$

is constant in time, while

$$\bar{x}^2 = \int x^2 \psi^* \psi dx = \hbar^2 \int \frac{\partial \psi^*}{\partial p} \cdot \frac{\partial \psi}{\partial p} dp$$

changes with time. Indeed, from the latter relation and eq. (3.2), we have

$$\bar{x}^2 = \hbar^2 \int \frac{\partial A^*}{\partial p} \cdot \frac{\partial A}{\partial p} dp + i \hbar t \int \frac{\partial E}{\partial p} \left(A^* \frac{\partial A}{\partial p} - A \frac{\partial A^*}{\partial p} \right) dp + t^2 \int \left(\frac{\partial E}{\partial p} \right)^2 A^* A dp \quad (3.28)$$

or

$$\bar{x}^2 = \hbar^2 \int \frac{\partial A^*}{\partial p} \cdot \frac{\partial A}{\partial p} dp + \frac{i \hbar t}{m} \int p \left(A^* \frac{\partial A}{\partial p} - A \frac{\partial A^*}{\partial p} \right) dp + \frac{t^2}{m^2} \bar{p}^2. \quad (3.28')$$

The mean cross-section of an arbitrary wave-packet along each co-ordinate direction is, therefore, a quadratic function of time for the free-particle case. It, therefore, increases rapidly both before and after passing through a minimum. It is easy to re-express (3.28') in co-ordinate space. If ψ_0 denotes the value of ψ for $t = 0$, $(\bar{x}^2)_0 = \int x^2 \psi_0^* \psi_0 d^3x$, the value of (\bar{x}^2) at time $t = 0$ and $\vec{i}_0 = \frac{\hbar}{2mi} (\psi_0^* \cdot \text{grad } \psi_0 - \psi_0 \cdot \text{grad } \psi_0^*)$, the value of \vec{i} at time $t = 0$, then we obtain

$$\bar{x}^2 = (\bar{x}^2)_0 + 2t \int (x \vec{i}_0) d^3x + \frac{t^2}{m^2} \bar{p}^2 \quad (3.29')$$

and in the unprimed co-ordinate system ($\rho = \psi^* \psi$)

$$\overline{\Delta x^2} = (\overline{\Delta x^2})_0 + 2t \int (x - \bar{x}) \left(i_0 - \frac{\rho_0}{m} p \right) d^3x + \frac{t^2}{m^2} \overline{(\Delta p^2)}. \quad (3.29)$$

This result is not specially characteristic of Quantum Mechanics since the same holds for a system of freely moving points which are distributed with density ρ and current density \vec{i} as well as with the mean square deviation of the momentum, $(\overline{\Delta p^2})$. It must, however, be remembered that the continuity of \bar{x} and $\overline{\Delta x^2}$ with respect to time for arbitrarily small $(\Delta x^2)_0$ is essential for the possibility of repeating the measurement of position. In the same way one can calculate the temporal change in the mean value, $\Delta x_i \Delta x_m$ of the product of two co-ordinates, in the three-dimensional case. We then obtain an expression again quadratic in t :

$$\begin{aligned} \overline{\Delta x_i \Delta x_m} &= (\overline{\Delta x_i \Delta x_m})_0 + t \int \left[(x_i - \bar{x}_i) \left(i_m - \frac{\rho_0}{m} p_m \right) + \right. \\ &\quad \left. + (x_m - \bar{x}_m) \left(i_i - \frac{\rho_0}{m} p_i \right) \right] d^3x + \frac{t^2}{m^2} \overline{\Delta p_i \Delta p_m}. \end{aligned} \quad (3.30)$$

The characteristic feature of Quantum Mechanics, however, is that there exists a relation between the values $(\Delta x)^2$ and $(\Delta p)^2$ corresponding to the uncertainty relation in the sense that both expressions cannot be made arbitrarily small simultaneously.⁷ We can understand this easily by rearranging the inequality

$$D = \left| \frac{x}{2x^2} \psi + \frac{\partial \psi}{\partial x} \right|^2 \geq 0.$$

Then

$$\begin{aligned} D &= \frac{x^2}{4(x^2)^2} \psi \psi^* + \frac{x}{2x^2} \left(\psi \frac{\partial \psi^*}{\partial x} + \psi^* \frac{\partial \psi}{\partial x} \right) + \frac{\partial \psi}{\partial x} \frac{\partial \psi^*}{\partial x} \\ &= \frac{1}{4} \left(\frac{x}{x^2} \right)^2 \psi \psi^* + \frac{1}{2} \frac{\partial}{\partial x} \left(\frac{x}{x^2} \psi \psi^* \right) - \frac{1}{2} \frac{1}{x^2} \psi \psi^* + \frac{\partial \psi}{\partial x} \frac{\partial \psi^*}{\partial x} \\ &= \frac{1}{4} \frac{1}{(x^2)^2} \cdot [x^2 - 2x^2] \psi \psi^* + \frac{1}{2} \frac{\partial}{\partial x} \left(\frac{x}{x^2} \psi \psi^* \right) + \frac{\partial \psi}{\partial x} \frac{\partial \psi^*}{\partial x}. \end{aligned}$$

Hence, on integrating we get

$$\int D d^3x = \frac{1}{\hbar^2} \overline{p^2} - \frac{1}{4} \frac{1}{\overline{x^4}} \geq 0.$$

Thus

$$\overline{p^2} \overline{x^4} = (\overline{\Delta p})^2 (\overline{\Delta x})^2 \geq \frac{\hbar^4}{4}. \quad (3.31)$$

This is the quantitative refinement of the uncertainty relation. The equality sign in (3.31) holds only if

$$\frac{1}{2} \frac{x}{x^2} \psi + \frac{\partial \psi}{\partial x} = 0$$

or

$$\psi = C e^{-\frac{1}{4} \frac{x^4}{x^2}}. \quad (3.32)$$

If we are interested in the product $(\overline{\Delta p_i^2})(\overline{\Delta x_i^2})$ only for a definite value of the index i , then it is immaterial what its dependence on the other co-ordinates is. If this product is to attain the minimum simultaneously in all three co-ordinates, we must set

$$\psi = C e^{-\frac{1}{4} \left(\frac{x_1^4}{x_1^2} + \frac{x_2^4}{x_2^2} + \frac{x_3^4}{x_3^2} \right)}. \quad (3.32')$$

While $\overline{\Delta p_i^2}$ is constant in time, $\overline{\Delta x_i^2}$ changes with time. If the minimum of $(\overline{\Delta p_i})^2 (\overline{\Delta x_i})^2$ is to be reached for $t = 0$, the term linear in t in (3.29) must vanish (which can also be verified directly). For earlier and later times, therefore, the product in question has a larger value. By subsequent measurements it can indeed be further reduced, but it can never go below the minimum.

On account of the complete symmetry of the minimum problem with respect to p_x and x , the momentum space wave function $\phi(p)$ which corresponds to the above

⁷ Cf. for this H. Weyl, Group Theory and Quantum Mechanics, 2 edn., Appendix 1, Leipzig (1931), (English translation) Dover Publications Inc., New York (1950); W. Heisenberg, The Physical Principles of Quantum Theory, loc. cit. For generalisations, E.U. Condon, Science, Lancaster, Pa. (1929); H.P. Robertson, Phys. Rev. 34, 163 (1929) and particularly E. Schrödinger, Berl. Ber. (1930), 296, where theorems of the form (3.28), (3.29) are proved in general for the first time.

minimum, is likewise the Gaussian error function

$$\varphi(\vec{p}) = C e^{-\frac{1}{4} \frac{\vec{p}_1^2}{(4\vec{p}_1)^2}} \quad (3.33)$$

or

$$\varphi(\vec{p}) = C e^{-\frac{1}{4} \left[\frac{\vec{p}_1^2}{(4\vec{p}_1)^2} + \frac{\vec{p}_2^2}{(4\vec{p}_2)^2} + \frac{\vec{p}_3^2}{(4\vec{p}_3)^2} \right]}. \quad (3.33')$$

This can also be verified by a direct calculation using eq. (3.4).

Finally, we describe a general method to solve the time-dependent eq. (3.6) if space wave function ψ_0 at time $t = 0$ is given. This problem can be immediately solved if we succeed in finding a "fundamental solution" $U(\vec{x}; t)$ with the property that U is singular at $t = 0$ in such a way that for every finite region of integration, we have

$$\lim_{t \rightarrow 0} \int_V U d^3x = \begin{cases} 1, & \text{if the origin lies in } V, \\ 0, & \text{if the origin lies outside } V. \end{cases} \quad (3.34)$$

Due to the linear nature of the differential equation, the desired solution is

$$\begin{aligned} \psi(x_i; t) &= - \int U(\vec{x}_i - \vec{x}_i; t) \psi(\vec{x}_i; 0) d^3\vec{x} \\ &= - \int U(\vec{x}_i; t) \psi(\vec{x}_i + \vec{x}_i; 0) d^3\vec{x}. \end{aligned} \quad (3.35)$$

In order to find the fundamental solution U for the free-particle case of non-relativistic quantum mechanics, it is convenient to remember the formal similarity between the differential equation

$$\frac{\partial \psi}{\partial t} = \frac{i\hbar}{2m} \Delta \psi \quad (3.36)$$

and heat conduction or diffusion equation.⁸ (We have put $E_0 = 0$ in (3.36), for simplicity, but it can as well be achieved by separating the factor $e^{-iE_0 t/\hbar}$ from the wave function.) Here, however, we have to replace the coefficient of heat conduction by an imaginary quantity. Our fundamental solution then corresponds to a solution of the equation for heat conduction which is associated with a "heat pole" and in the one-dimensional case is of the form

$$U(x, t) = \frac{C}{\sqrt{t}} e^{+\frac{i\pi m}{2\hbar} \frac{x^2}{t}}.$$

It can be verified that this function satisfies the differential equation

$$\frac{\partial \psi}{\partial t} = \frac{i\hbar}{2m} \frac{\partial^2 \psi}{\partial x^2}.$$

Then

$$\int_{x_1}^{x_2} U(x, t) dx = C \sqrt{\frac{2\hbar}{m}} \int_{\frac{\sqrt{\frac{m}{2\hbar}} \frac{x_1}{t}}{\sqrt{t}}}^{\frac{\sqrt{\frac{m}{2\hbar}} \frac{x_2}{t}}{\sqrt{t}}} e^{i\pi n} d\xi.$$

⁸ For this similarity, consult particularly P. Ehrenfest, Z. Physik 45, 455 (1927); cf. for the following also L. de Broglie, Wave Mechanics, Chapter 13.

Now as $a \rightarrow +\infty$ and $b \rightarrow +\infty$,

$$\lim_{a \rightarrow +\infty} \int_a^b e^{i k \xi} d\xi = 0.$$

But as $a \rightarrow -\infty$ and $b \rightarrow +\infty$, we obtain

$$\lim_{a \rightarrow -\infty} \int_a^b e^{i k \xi} d\xi = \int_{-\infty}^{\infty} e^{i k \xi} d\xi = \sqrt{\pi} e^{i \pi/4}.$$

Therefore, we have, as required by (3.34):

$$\lim_{t \rightarrow 0} \int_{x_1}^{x_2} U dx = \begin{cases} 0, & \text{if the origin } x=0 \text{ lies outside the interval } (x_1, x_2), \\ 1, & \text{inside} \end{cases}$$

normalising the constant C to

$$C = e^{-i \pi/4} \sqrt{\frac{m}{2 \pi \hbar}}.$$

Finally

$$U(x, t) = e^{-i \pi/4} \sqrt{\frac{m}{2 \pi \hbar}} \frac{1}{\sqrt{t}} e^{\frac{im}{2\hbar} \frac{x^2}{t}}. \quad (3.37)$$

From this we obtain immediately in the three-dimensional case

$$U(x_1, x_2, x_3, t) = U(x_1, t) U(x_2, t) U(x_3, t) = e^{-\frac{3\pi}{4} i} \left(\frac{m}{2\pi\hbar}\right)^{\frac{3}{2}} t^{-\frac{3}{2}} e^{\frac{im}{2\hbar} \frac{x_1^2+x_2^2+x_3^2}{t}}. \quad (3.38)$$

Substituting this expression into (3.35), we get the general solution $\psi(x_1, x_2, x_3, t)$ of the wave equation.⁹ To find the solution U one could have also started from its decomposition into spatial Fourier components according to (3.1) and (3.4) and on using (3.34).

4. The Wave Function of a Particle Acted on by Forces

The description of the states of a system of particles subject to forces by statistical laws emerged as a generalisation of that for free particles. Evidently, the corresponding concepts and laws should be consistent with one another and contain the laws of classical particle mechanics as limiting cases. Apart from these general requirements, only the results can justify the need for particular assumptions. Neglecting the relativistic corrections we can formulate these assumptions for the single particle case in the following manner:

1. At a fixed time t , the probability of finding the position co-ordinates x_i of the particle between x_i and $x_i + dx_i$ is a well-defined concept. It is again given by

$$W(x_1 x_2 x_3; t) dx = \psi^* \psi dx, \quad (4.1)$$

⁹ Particular solutions of the wave equation, especially for the case in which $\psi(x, ; 0)$ is given by the Gaussian error function (3.32), can be found in W. Heisenberg, Z. Physik 43, 172 (1927); E.H. Kennard, Z. Physik 44, 326 (1927); C.G. Darwin, Proc. Roy. Soc. Lond., Ser. A 117, 258 (1927).

where $\psi(\vec{x}, t)$ is in general a complex wave function which is not by itself observable; ψ^* denotes the complex conjugate of ψ , and ψ is normalised according to

$$\int \psi \psi^* d^3x = 1. \quad (4.2)$$

This assumption is natural since the measurement of the position can take place in such a short time that the presence of the forces has no role to play. From (4.2) we get the condition

$$\frac{d}{dt} \int \psi \psi^* d^3x = 0$$

for the variation of ψ with time. This condition can be satisfied at each instant of time only if $\partial\psi/\partial t$ and $\partial\psi^*/\partial t$ are determined, once ψ and ψ^* are given. (Regarding the need for functions with many components in the case of particles with spin, cf. Sec. 13.)

2. If we set

$$-\frac{\hbar}{i} \frac{\partial\psi}{\partial t} = H\psi, \quad (4.3)$$

then H cannot be a general operator, but should be a *linear* operator. As already mentioned above in § 3, eq. (3.9) this requirement associates a new function $H\psi$ with the function ψ , which has the property that for arbitrary constant c , which can be complex, we have

$$H(c\psi) = cH\psi,$$

and for two arbitrary functions, ψ_1, ψ_2

$$H(\psi_1 + \psi_2) = H\psi_1 + H\psi_2.$$

[From these two properties we see that $H\psi$ does not explicitly depend on ψ^* .]

The condition that the operator H be linear can be taken as a generalisation of the *superposition principle*, since in the free-particle case (as we have already seen) it directly expresses the principle stemming from wave theory. This principle is essential for a consistent formulation of the notion of measurement when the coupling of the system with the measuring apparatus is itself described quantum-mechanically (cf. Sec. 9).

In order that the constancy of $\int \psi \psi^* d^3x$ follows from eq. (4.3), the operator H must have the property

$$\int [\psi^* H\psi - \psi (H\psi)^*] d^3x = 0. \quad (4.4)$$

Here use has been made of the wave equation

$$+\frac{\hbar}{i} \frac{\partial\psi^*}{\partial t} = (H\psi)^*. \quad (4.3^*)$$

This must hold for all regular functions, which vanish sufficiently fast at infinity. An operator H , possessing these properties is called *Hermitian*.¹ On account of the

¹ It is to be noted here that from (4.4) alone the linearity of H does not follow. For example the non-linear operator $H\psi = i\psi \cdot \partial\psi^*/\partial x$ (where $(H\psi)^* = -i\psi^* \cdot \partial\psi/\partial x$) also has the property (4.4), since $\psi^* H\psi - \psi (H\psi)^* = \frac{i}{2} \frac{\partial}{\partial x} (\psi^2 \psi^*)$. It is, therefore, necessary to introduce the superposition principle as a new assumption.

linearity of H , it follows from (4.4) for two arbitrary functions that

$$\int [\psi_1^* H \psi_2 - \psi_2^* (H \psi_1)^*] d^3x = 0. \quad (4.4')$$

3. The relations (3.1'') and (3.4) connecting ψ with the "amplitude" $\varphi(\vec{p})$ which determines the probability $W(p, t)d^3p$ for the momentum in conformity with (3.16) are retained here as well;² only, this probability is no longer a constant now. (On the measurement of momentum for bound particles, cf. Sec. 15.) The conditions (3.23) and (3.24) as also the completeness relation (3.5) remain true.

From the standpoint of non-relativistic wave mechanics, the only way of determining the operator H for a particular system is by comparing the behaviour of the general solution of eq. (4.3) in suitable limiting cases as given by Bohr's correspondence principle, with the properties of the trajectories of the same system according to classical mechanics. Only experience with different examples can decide between the various choices for H arising from the correspondence principle.

As a simple example we consider a particle in an external force-field with the potential function $V(\vec{x})$. For this case, the classical Hamiltonian is

$$H(\vec{p}, \vec{x}) = \sum_i \frac{\vec{p}_i^2}{2m} + V(\vec{x}).$$

Remembering that the expectation value of p_i^2 , according to (3.24') amounts to

$$\bar{p}_i^2 = \int p_i^2 |\varphi(\vec{p})|^2 d^3p = -\hbar^2 \int \psi^* \frac{\partial^2 \psi}{\partial x_i^2} d^3x$$

with Schrödinger,³ we assume the wave equation to have the obvious form

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi + V \psi. \quad (4.5)$$

We shall draw some conclusions about the mean values, which are analogous to the rules regarding the behaviour of the mid-points and the cross-section of the wave-packets, formulated in the foregoing section. From (4.5) follows again the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{i} = 0$$

with $\rho = \psi^* \psi$ and with the original expression (3.15)

$$i_k = \frac{\hbar}{2m i} \left(\psi^* \frac{\partial \psi}{\partial x_k} - \psi \frac{\partial \psi^*}{\partial x_k} \right)$$

for the current density; for, in the calculation of $\partial \rho / \partial t$, the term with $V \psi$ drops out.

From this follow the relations given by (3.17), (3.22), and (3.23)

$$\begin{aligned} \bar{x}_k &= \int x_k \psi^* \psi d^3x; \quad \bar{p}_k = \int p_k \psi^* \psi d^3p = \int \psi^* \left(\frac{i}{\hbar} \frac{\partial \psi}{\partial x_k} \right) d^3x; \\ \frac{d\bar{x}_k}{dt} &= \int i_k d^3x = \frac{1}{m} \bar{p}_k = \left(\frac{\partial H}{\partial p_k} \right). \end{aligned}$$

²This was noticed in a general context by P. Jordan, Z. Physik **40**, 809 (1927).

³E. Schrödinger, Ann. d. Phys. **79**, 361 (1926). M. Born [Z. Physik **38**, 803 (1926)], in particular, pointed out in his treatment of the collision process, the necessity for a statistical interpretation of the wave function.

We also obtain something new, when we calculate the rate of change $d\bar{p}_k/dt$ of \bar{p}_k with time. For the free-particle case this derivative vanishes but it is no longer so in the present case. To show this, let us construct

$$\begin{aligned} m \frac{\partial i_k}{\partial t} &= \frac{1}{2} \left[(\mathbf{H}\psi)^* \frac{\partial \psi}{\partial x_k} - \psi^* \frac{\partial}{\partial x_k} (\mathbf{H}\psi) + (\mathbf{H}\psi) \frac{\partial \psi^*}{\partial x_k} - \psi \frac{\partial}{\partial x_k} (\mathbf{H}\psi)^* \right] \\ &= \frac{\hbar^2}{4m} \left[-(\Delta\psi^*) \frac{\partial \psi}{\partial x_k} + \psi^* \frac{\partial}{\partial x_k} (\Delta\psi) - (\Delta\psi) \frac{\partial \psi^*}{\partial x_k} + \psi \frac{\partial}{\partial x_k} (\Delta\psi^*) \right] \\ &\quad + \frac{1}{2} \left[V\psi^* \frac{\partial \psi}{\partial x_k} - \psi^* \frac{\partial}{\partial x_k} (V\psi) + V\psi \frac{\partial \psi^*}{\partial x_k} - \psi \frac{\partial}{\partial x_k} (V\psi^*) \right]. \end{aligned}$$

The second bracket simplifies immediately to $-\frac{\partial V}{\partial x_k} \psi^* \psi$. The first bracket is to be rearranged as follows. For arbitrary functions u, v , we have

$$v \Delta u - u \Delta v = \sum_i \frac{\partial}{\partial x_i} \left(v \frac{\partial u}{\partial x_i} - u \frac{\partial v}{\partial x_i} \right).$$

If we set in this $v = \psi^*$, $u = \frac{\partial \psi}{\partial x_k}$, first and then $v = \psi$, $u = \frac{\partial \psi^*}{\partial x_k}$ we obtain on introducing the force $K_i = -\frac{\partial V}{\partial x_i} = -\frac{\partial H}{\partial x_i}$

$$m \frac{\partial i_k}{\partial t} = - \sum_i \frac{\partial T_{ki}}{\partial x_i} + K_k \psi^* \psi \quad (4.6)$$

with

$$T_{ki} = \frac{\hbar^2}{4m} \left[-\psi^* \frac{\partial^2 \psi}{\partial x_k \partial x_i} - \psi \frac{\partial^2 \psi^*}{\partial x_k \partial x_i} + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi^*}{\partial x_i} + \frac{\partial \psi^*}{\partial x_k} \frac{\partial \psi}{\partial x_i} \right]. \quad (4.7)$$

The tensor T_{ki} , which also satisfies the symmetry condition

$$T_{ki} = T_{ik} \quad (4.7')$$

is called the stress tensor.⁴ One obtains from eq. (4.6)

$$\frac{d\bar{p}_k}{dt} = m \frac{d^3 \bar{x}_k}{dt^3} = m \int \frac{\partial i_k}{\partial t} d^3x = \int K_k \psi^* \psi d^3x = K_k = -\overline{\left(\frac{\partial V}{\partial x_k} \right)} = -\overline{\left(\frac{\partial H}{\partial x_k} \right)}, \quad (4.8)$$

which signifies that the time derivative of the mean value of p_k is equal to the mean value of the force taken over the wave-packet.⁵ The latter is in general different from the value of the force at the mid-point \bar{x}_k of the wave-packet. Only when the wave-packet can be chosen in conformity with the uncertainty relations $\Delta p_k \Delta x_k \sim \hbar$, so that the force varies slowly within the packet, do we obtain a packet which behaves like a classical particle with a trajectory satisfying the equation of motion (cf. Sec. 12)

$$m \frac{d^3 \bar{x}_k}{dt^3} = -\frac{\partial V}{\partial x_k}.$$

⁴ A relativistic generalisation of this is to be found in E. Schrödinger, Ann. d. Phys. 82, 265 (1927); cf. also Sec. 21, Chapter IX.

⁵ P. Ehrenfest, Z. Physik 45, 455 (1927).

Another consequence of eq. (4.6) concerns the virial theorem.⁶ Multiplying (4.6) by x_k and integrating by parts we have

$$m \frac{d}{dt} \int x_k i_k d^3x = + \int T_{kk} d^3x - \int x_k \frac{\partial V}{\partial x_k} \psi^* \psi d^3x.$$

On account of (4.7), we get for the first integral on the right-hand side, after partial integration

$$\frac{\hbar^2}{m} \int \frac{\partial \psi^*}{\partial x_k} \frac{\partial \psi}{\partial x_k} d^3x = \frac{\overline{p_k^2}}{m}.$$

Hence, the above equation takes the form

$$m \frac{d}{dt} \int x_k i_k d^3x = \frac{\overline{p_k^2}}{m} - \left(\overline{x_k \frac{\partial V}{\partial x_k}} \right). \quad (4.9)$$

Summing over the index k , we obtain the analogue of the virial theorem.

Finally, one can consider, as in (3.29), the temporal change in the cross-section of a wave-packet, given by

$$\overline{(\Delta x_k)^2} = \int (x_k - \overline{x_k})^2 \psi^* \psi d^3x. \quad (4.10)$$

But it will in general no longer be possible on account of the presence of forces to give the variation of $\overline{(\Delta x_k)^2}$ for finite times; instead of this we can calculate the infinitesimal changes given by the first and second differential coefficients of $\overline{(\Delta x_k)^2}$ with respect to time. Since $\int (x_k - \overline{x_k}) \psi^* \psi d^3x = 0$, we get immediately

$$\frac{d}{dt} \overline{(\Delta x_k)^2} = \int (x_k - \overline{x_k})^2 \frac{\partial}{\partial t} (\psi^* \psi) d^3x.$$

On using the continuity equation and integrating by parts we get

$$\frac{d}{dt} \overline{(\Delta x_k)^2} = 2 \int (x_k - \overline{x_k}) i_k d^3x. \quad (4.11)$$

Because of

$$\frac{d \overline{x_k}}{dt} = \int i_k d^3x = \frac{\overline{p_k}}{m}$$

we further have

$$\frac{1}{2} \frac{d^2}{dt^2} \overline{(\Delta x_k)^2} = \int (x_k - \overline{x_k}) \frac{\partial i_k}{\partial t} d^3x - \left(\int i_k d^3x \right)^2$$

and on using (4.9)

$$\frac{m}{2} \frac{d^2}{dt^2} \overline{(\Delta x_k)^2} = \frac{\overline{p_k^2} - \overline{(\overline{p_k})^2}}{m} + \overline{(\Delta x_k \Delta K_k)}$$

or

$$\frac{m}{2} \frac{d^2}{dt^2} \overline{(\Delta x_k)^2} = \frac{\overline{(p - \overline{p_k})^2}}{m} + \overline{(\Delta x_k \Delta K_k)}. \quad (4.12)$$

⁶ A. Sommerfeld, Wave Mechanics, Dutton, New York (1929).

The relations (4.11) and (4.12) represent the natural generalisation of (3.29).

Before we take up the question of the interaction between many particles, we shall give the modifications necessary in the wave equation, in the presence of an external magnetic field. If Φ_k are the components of the vector potential, e the charge of the particle and c the velocity of light, the magnetic field strength is given by⁷

$$\mathcal{H}_{kl} = \frac{\partial \Phi_l}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_l}. \quad (4.13)$$

The electric field strength has an additional term

$$\mathcal{E}_k = -\frac{1}{c} \frac{\partial \Phi_k}{\partial t}, \quad (4.14)$$

if Φ_k depends explicitly on the time. The force on the particle will then be

$$\left. \begin{aligned} K_k &= -\frac{\partial V}{\partial x_k} + e \left(\mathcal{E}_k + \frac{1}{c} \sum_l \mathcal{H}_{kl} \dot{x}_l \right) \\ &= -\frac{\partial V}{\partial x_k} + \frac{e}{c} \left[-\frac{\partial \Phi_k}{\partial t} + \sum_l \left(\frac{\partial \Phi_l}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_l} \right) \dot{x}_l \right]. \end{aligned} \right\} \quad (4.15)$$

As is well known (with this expression for the force), the equations of motion

$$m \frac{d^2 x_k}{dt^2} = K_k$$

can be written in the canonical form⁸

$$\frac{dx_k}{dt} = \frac{\partial H}{\partial p_k}, \quad \frac{dp_k}{dt} = -\frac{\partial H}{\partial x_k},$$

if the operator

$$H = \sum_k \frac{1}{2m} \left(p_k - \frac{e}{c} \Phi_k \right)^2 + V(\vec{x}) \quad (4.16)$$

Then

$$\dot{x} = \frac{1}{m} \left(p_k - \frac{e}{c} \Phi_k \right); \quad p_k = m \dot{x}_k + \frac{e}{c} \Phi_k; \quad (4.16')$$

thus the relationship between momentum and velocity is altered.

The fact that the classical Hamiltonian (4.16) is obtained from the one without the magnetic field by replacing p_k by $p_k - \frac{e}{c} \Phi_k$ makes it reasonable to write the wave equation of the particle in a magnetic field by replacing the operator $\frac{\hbar}{i} \frac{\partial}{\partial x_k}$ by $\frac{\hbar}{i} \frac{\partial}{\partial x_k} - \frac{e}{c} \Phi_k$ in the wave equation (4.5) without the magnetic field. We then

⁷ We prefer to write \mathcal{H} as an antisymmetric tensor ($\mathcal{H}_{kl} = -\mathcal{H}_{lk}$) so that $\mathcal{H}_{23}, \mathcal{H}_{31}, \mathcal{H}_{12}$ denote respectively the 1, 2, 3-components of \mathcal{H} . The vector product $[\vec{x}, \vec{\mathcal{H}}]$ then has the 1-component $\dot{x}_2 \mathcal{H}_{12} - \dot{x}_3 \mathcal{H}_{21}$ and this is, in fact, equal to $\sum_l \mathcal{H}_{ll} \dot{x}_l$ since $\mathcal{H}_{31} = -\mathcal{H}_{13}, \mathcal{H}_{11} = 0$.

⁸ Historically speaking, it may be remarked that this was first shown by Larmor in his book, Aether and Matter, Cambridge (1900).

obtain, in place of (4.5), the general equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \frac{1}{2m} \sum_k \left(\frac{\hbar}{i} \frac{\partial}{\partial x_k} - \frac{e}{c} \Phi_k \right) \left(\frac{\hbar}{i} \frac{\partial}{\partial x_k} - \frac{e}{c} \Phi_k \right) \psi + V \psi = 0, \quad (4.17)$$

which can also be written as

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \frac{1}{2m} \sum_k \left[-\hbar^2 \frac{\partial^2}{\partial x_k^2} \psi - \frac{\hbar e}{ic} \frac{\partial}{\partial x_k} (\Phi_k \psi) - \frac{\hbar e}{ic} \Phi_k \frac{\partial \psi}{\partial x_k} + \frac{e^2}{c^2} \Phi_k^2 \psi \right] + V \psi = 0. \quad (4.17')$$

The justification for the above ansatz is that from eq. (4.17') the rules regarding the mean values of p_k and x_k and the total current $i_k = \int i_k d^3x$ and their time derivatives follow, which are analogous to the corresponding equations of motion of classical mechanics.

We have similar to (3.14) again a continuity equation

$$\frac{\partial}{\partial t} (\psi^* \psi) + \operatorname{div} \vec{i} = 0,$$

which at the same time verifies that H is Hermitian. For the current, however, we have the modified expression

$$i_k = \frac{1}{2m} \left[\psi^* \left(\frac{\hbar}{i} \frac{\partial}{\partial x_k} - \frac{e}{c} \Phi_k \right) \psi - \psi \left(\frac{\hbar}{i} \frac{\partial}{\partial x_k} + \frac{e}{c} \Phi_k \right) \psi^* \right] \quad (4.18)$$

or

$$i_k = \frac{\hbar}{2mi} \left(\psi^* \frac{\partial \psi}{\partial x_k} - \psi \frac{\partial \psi^*}{\partial x_k} \right) - \frac{e}{mc} \Phi_k \psi^* \psi. \quad (4.18')$$

Forming

$$\bar{p}_k = \int p_k \varphi^* \varphi d^3p = \int \psi^* \frac{\hbar}{i} \frac{\partial \psi}{\partial x_k} d^3x$$

and

$$\frac{d\bar{x}_k}{dt} = \frac{d}{dt} \int x_k \psi^* \psi d^3x = \int i_k d^3x, \quad (3.22')$$

(which corresponds to eq. 3.22) we find

$$\frac{d\bar{x}_k}{dt} = \frac{1}{m} \left(\bar{p}_k - \frac{e}{c} \bar{\Phi}_k \right), \quad (4.16'')$$

which is similar to (4.16').

Further we find, analogous to (4.6) and (4.7)

$$m \frac{\partial i_k}{\partial t} = - \sum_l \frac{\partial T_{kl}}{\partial x_l} + \left(-\frac{\partial V}{\partial x_k} - \frac{e}{c} \frac{\partial \Phi_k}{\partial t} \right) \psi^* \psi + \frac{e}{c} \sum_l \mathcal{H}_{kl} i_l \quad (4.19)$$

where

$$T_{kl} = \frac{\hbar^2}{4m} \left\{ -\psi^* \left(\frac{\partial}{\partial x_l} - \frac{ie}{hc} \Phi_l \right) \left(\frac{\partial \psi}{\partial x_k} - \frac{ie}{hc} \Phi_k \psi \right) - \psi \left(\frac{\partial}{\partial x_l} + \frac{ie}{hc} \Phi_l \right) \left(\frac{\partial \psi^*}{\partial x_k} + \frac{ie}{hc} \Phi_k \psi^* \right) + \left(\frac{\partial \psi}{\partial x_k} - \frac{ie}{hc} \Phi_k \psi \right) \left(\frac{\partial \psi^*}{\partial x_l} + \frac{ie}{hc} \Phi_l \psi^* \right) + \left(\frac{\partial \psi^*}{\partial x_k} + \frac{ie}{hc} \Phi_k \psi^* \right) \left(\frac{\partial \psi}{\partial x_l} - \frac{ie}{hc} \Phi_l \psi \right) \right\} \quad (4.20)$$

or

$$T_{kl} = \frac{\hbar^2}{4m} \left\{ \left[-\psi^* \frac{\partial^2 \psi}{\partial x_l \partial x_k} - \psi \frac{\partial^2 \psi^*}{\partial x_l \partial x_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial \psi^*}{\partial x_k} + \frac{\partial \psi^*}{\partial x_l} \frac{\partial \psi}{\partial x_k} \right] + \frac{2ie}{hc} \left[\Phi_k \left(\psi^* \frac{\partial \psi}{\partial x_l} - \psi \frac{\partial \psi^*}{\partial x_l} \right) + \Phi_l \left(\psi^* \frac{\partial \psi}{\partial x_k} - \psi \frac{\partial \psi^*}{\partial x_k} \right) \right] + \frac{4e^2}{hc^2} \Phi_k \Phi_l \psi^* \psi \right\}, \quad (4.21)$$

so that the symmetry condition $T_{kl} = T_{lk}$ is again fulfilled. Remembering (4.15) and setting

$$\bar{K}_k = \int \left[-\left(\frac{\partial V}{\partial x_k} + \frac{e}{c} \frac{\partial \Phi_k}{\partial t} \right) \psi^* \psi + \frac{e}{c} \sum_l \mathcal{H}_{kl} i_l \right] d^3x, \quad (4.22)$$

we obtain from (4.19) since $\int i_k d^3x = \frac{d \bar{x}_k}{dt}$,

$$m \frac{d^2 \bar{x}_k}{dt^2} = \bar{K}_k \quad (4.23)$$

as the analogue of the equation of motion. Further with

$$\bar{x}_k \bar{K}_k = \int \left[-x_k \left(\frac{\partial V}{\partial x_k} + \frac{e}{c} \frac{\partial \Phi_k}{\partial t} \right) \psi^* \psi + \frac{e}{c} \sum_l \mathcal{H}_{kl} x_k i_l \right] d^3x \quad (4.22')$$

we have exactly as before

$$m \frac{d}{dt} \int x_k i_k d^3x = \int T_{kk} d^3x + \bar{x}_k \bar{K}_k.$$

By integration by parts, it follows from (4.20) that

$$\begin{aligned} \int T_{kk} d^3x &= \frac{\hbar^2}{m} \int \left(\frac{\partial \psi^*}{\partial x_k} + \frac{i e}{\hbar c} \Phi_k \psi^* \right) \left(\frac{\partial \psi}{\partial x_k} - \frac{i e}{\hbar c} \Phi_k \psi \right) d^3x \\ &= -\frac{1}{m} \overline{\left(p_k - \frac{e}{c} \Phi_k \right)^2} = m \bar{\dot{x}}_k^2. \end{aligned}$$

The last two expressions can indeed be justified only by using a systematic operator calculus; we shall speak about it later. With this proviso, we obtain

$$m \frac{d}{dt} \int x_k i_k d^3x = m \bar{\dot{x}}_k^2 + \bar{x}_k \bar{K}_k, \quad (4.24)$$

which is the analogue of the virial theorem (4.9). Similar to (4.11) and (4.12), we have

$$\frac{d}{dt} \overline{(\Delta x_k)^2} = 2 \int (x_k - \bar{x}_k) i_k d^3x, \quad (4.25)$$

$$\frac{m}{2} \frac{d^2}{dt^2} \overline{(\Delta x_k)^2} = m \overline{(\dot{x}_k - \bar{\dot{x}}_k)^2} + \overline{(x_k - \bar{x}_k) K_k}. \quad (4.26)$$

It is a well-known and important fact that the potentials Φ_k are determined only up to an additional gradient term, since by such an addition the magnetic field strengths \mathcal{H}_{kl} are not altered. Hence it is permitted to make the substitution

$$\Phi'_k = \Phi_k + \frac{\partial f}{\partial x_k}, \quad (4.27)$$

where f is an arbitrary function of the position co-ordinates. Since f can contain the time explicitly, one has to set, at the same time,

$$V' = V - \frac{e}{c} \frac{\partial f}{\partial t}, \quad (4.27')$$

in order to keep the expression (4.15) for the force, invariant. In fact, we then have

$$\frac{\partial V'}{\partial x_k} + \frac{e}{c} \frac{\partial \Phi'_k}{\partial t} = \frac{\partial V}{\partial x_k} + \frac{e}{c} \frac{\partial \Phi_k}{\partial t}; \quad \mathcal{H}_{kl} = \mathcal{H}'_{kl}.$$

Since in the wave equation (4.17) not only the magnetic and electric field strengths but also the potentials V and Φ_k enter, it may at first sight appear as if the physical results following from this wave equation also depend on the absolute values of the potentials. This is, however, not so; if ψ is a solution of the wave equation (4.17) with the potentials V and Φ_k , we obtain a solution ψ' , with the potentials V' and Φ'_k given by (4.27) and (4.27') if we make the substitution

$$\psi' = \psi \exp\left(\frac{ie}{\hbar c} t\right). \quad (4.27'')$$

The group of substitutions defined by (4.27), (4.27'), and (4.27'') is called the *gauge group*. Quantities which do not change under these substitutions are called *gauge-invariant quantities*.⁹ It is worth noting that not only the probability density $\psi\psi^*$, but also the current \vec{i} given by (4.18) and the stress tensor T_{kl} , defined by (4.20), are gauge-invariant quantities. From this point of view the form of the wave equation (4.17) and in particular the special choice of the Hamiltonian operator in this equation, must be considered as quite natural. On the other hand, this equation depends essentially on the assumption that the field quantities V and Φ_k themselves can be considered as classical quantities (given space-time functions) of such a kind that the possible influence of the quantum of action on the definition of these field quantities can be disregarded.

5. Many-Particle Interactions – Operator Calculus

The manner in which a composite system consisting of many sub-systems is described in quantum mechanics is of fundamental importance and is characteristic of the theory. It shows, on the one hand, the fruitfulness of the idea of Schrödinger in introducing the ψ -function, which satisfies a linear equation and, on the other, the purely symbolic nature of his function that is quite different from the wave functions of the classical theory (surface waves of liquids, elastic waves, electromagnetic waves, etc.).

When a system of many particles is given, we do *not* obtain a sufficient description of the system by a statement about the probability of finding *one* of the particles at a definite position. Let us consider, for example, a system consisting of two material particles which are placed inside a closed box. Let this box be divided into two parts by a wall with a small opening which can be closed. By closing the opening suddenly and taking the two halves apart, one can determine in which half of the box each of the two particles is present at the instant considered. We can now investigate not only how high the probability is for each particle to be found in one or

⁹ The invariance of the wave equation under the group of substitutions in question was (in the case of a relativistic generalisation of this equation) first given by V. Fock, Z. Physik 39, 226 (1927). The analogy of this group to the gauge group in an earlier theory of Weyl on gravitation and electricity was brought out by F. London, Z. Physik 42, 375 (1927). The connection between this group and the conservation law for charge was established by Weyl himself while deriving the wave equation from a variational principle [Z. Physik 56, 330 (1929)]. About the gauge group in the relativistic wave equation cf. Sec. 21.

the other half of the box, but how frequently the particles can be found in the same half or in different halves of the box. Instead of the separating wall, we can also use a "microscope" together with radiation of short wave-length and instead of dividing the finite volume (the box) into only two parts, we can make a finer sub-division of the space. Let there be N particles present and let $x_k^{[1]}, x_k^{[2]}, \dots, x_k^{[N]}$ be their coordinates which we simply write as $q_1 \dots q_f$, where $f = 3N$ denotes the number of degrees of freedom of the system. Further we shall simply write dq for the multi-dimensional volume element $dq_1 dq_2 \dots dq_f$. Thus the basic assumptions for the description of a system with several particles can be formulated in the following manner:

1. At each instant of time t there exists the probability

$$W(q_1 \dots q_f; t) dq \quad (5.1)$$

for finding simultaneously the co-ordinates of the first particle in the region $(q_k, q_k + dq_k)$ where $k = 1, 2, 3$, those of the second particle in $(q_k, q_k + dq_k)$ where $k = 4, 5, 6, \dots$, those of the N^{th} particle in $(q_k, q_k + dq_k)$ where $k = f - 2, f - 1, f$.

By way of explanation of this probability concept, it is to be remarked that here we have assumed the distinguishability of the particles; the probability of finding the first particle in the region $x_k^{[1]}, x_k^{[1]} + dx_k^{[1]}$ and the second in the region $x_k^{[2]}, x_k^{[2]} + dx_k^{[2]}$ will in general be different from the probability of finding the second particle in the region $x_k^{[1]}, x_k^{[1]} + dx_k^{[1]}$ and the first particle in the region $x_k^{[2]}, x_k^{[2]} + dx_k^{[2]}$. In other words, the probability depends on the order in which the $x_k^{[p]}$ appear as arguments $q_1 \dots q_f$ of W . Such distinguishability is surely possible if both the particles are of different types, e.g., when they have different masses (such as electron and proton or the nuclei of two different isotopes). The existence of identical entities in nature, such as, e.g., two electrons or two protons or two α -particles, demands particular care which, incidentally, still fails to show in the foundations of quantum mechanics developed till now. For particles of the same kind one can only ask for the probability that one particle is to be found in $(x_k^{[1]}, x_k^{[1]} + dx_k^{[1]})$, another in $(x_k^{[2]}, x_k^{[2]} + dx_k^{[2]})$ and yet another in $(x_k^{[N]}, x_k^{[N]} + dx_k^{[N]})$. If, therefore, many particles of the same kind are present, it is meaningful to speak of functions W which are symmetric in the co-ordinates of the identical particles. We shall return to the discussion of this case in detail in Sec. 14; for the time being we ignore it.

Integrating W over the co-ordinates of all particles except one, we obtain N new density functions

$$W_1(x_1, x_2, x_3), \quad W_2(x_4, x_5, x_6), \quad \dots \quad W_N(x_{3N-2}, x_{3N-1}, x_{3N}),$$

which give the probability that a particular particle is to be found in a particular place; here we do not raise the question as to where the other particles are to be found. These functions W_i ($i = 1, \dots, N$) really convey much less information about the system than the original function $W(q_1, \dots, q_f; t)$ of f space co-ordinates, in that the latter obviously cannot follow uniquely from the former, but we can certainly obtain the functions W_i from the function W . [In the above example of the box, divided into two halves and with two particles inside the box we cannot conclude anything about the relative frequencies, for example, of the cases: (1) "both the particles are in same half" and (2) "the two particles are in different halves" from the statement

"each of the particles has equal probability of being found in one or the other of the two halves of the box."]

Only in the special case, when the function W can be factorised in the form

$$W(q_1 \dots q_s) = W_1(q_1, q_2, q_3) W_2(q_4, q_5, q_6) \dots W_N(q_{3N-2}, q_{3N-1}, q_{3N})$$

is a knowledge of the functions W_1, \dots, W_N equivalent to a knowledge of the function $W(q_1, \dots, q_s)$. In this special case, we say that the particles are statistically independent of one another.

The existence of the probability function $W(q_1, \dots, q_s; t)$ implies the assumption that the measurement of position of a particular particle does not basically disturb the position measurement of other particles so that information regarding the position of a particle which may be made use of for predicting other measurements (e.g., the position of the same particle at a later time) is not lost by a knowledge of the position of the other particles. This state of affairs is closely connected with the question as to how far the *simultaneity* of the position measurements of the different particles is essential for the existence of the probability. That is to say: Under what circumstances does there exist a probability

$$W(x_k^{(1)}, t^{(1)}; x_k^{(2)}, t^{(2)}; \dots; x_k^{(N)}, t^{(N)}) dq_1 \dots dq_{3N} \quad (5.2)$$

that the first particle is found at time $t^{(1)}$ in the region $x_k^{(1)}, x_k^{(1)} + dx_k^{(1)}$, the second particle at time $t^{(2)}$ in the region $x_k^{(2)}, x_k^{(2)} + dx_k^{(2)}$, ... and the N^{th} particle at time $t^{(N)}$ in the region $x_k^{(N)}, x_k^{(N)} + dx_k^{(N)}$, respectively. In general, i.e. if any interaction is present between the particles, the absence of mutual disturbance of the measurements is guaranteed if and only if for the distance r_{ab} between any pair (a, b) of particles and the corresponding times we have the inequality

$$|t_a - t_b| < \frac{r_{ab}}{c}. \quad (5.3)$$

For, the change in the force exerted by particle a on particle b which arises from the position measurement of a , can propagate at most with the velocity of light c . If in relativistic quantum mechanics the concept of probability $W_1(x_1, x_2, x_3; t) dx_1 dx_2 dx_3$ for a particle to be at a given point at time t is meaningful, then the existence of the probability (5.2) requires that its arguments satisfy the condition (5.3).¹ In non-relativistic wave mechanics it is customary to consider c as infinitely large and then we restrict ourselves to the case, $t^{(1)} = t^{(2)} = \dots = t^{(N)} = t$.

2. As a straight-forward generalisation of the corresponding assumption in the case of a single particle, we assume the existence of a function²

$$\psi(q_1 \dots q_s; t)$$

such that

$$W(q_1 \dots q_s; t) dq = \psi^* \psi dq. \quad (5.4)$$

This function ψ must again satisfy an equation of the type (4.3):

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \mathbf{H} \psi$$

¹ See for this: P.A.M. Dirac, V. Fock and B. Podolsky, Phys. Z. Sowjet, **2**, 468 (1932); F. Bloch, Phys. Z. Sowjet, **5**, 301 (1934).

² About the necessity of many-component ψ -functions for particles with spin, cf. Sec. 13.

where H is a linear operator. The function $(H\psi)(q_1 \dots q_f; t)$ is thereby uniquely determined by the function $\psi(q_1 \dots q_f; t)$ defined at the same instant of time t , without our having to know ψ at other times. In order to satisfy the equation

$$\frac{d}{dt} \int \psi \psi^* dq = 0$$

H must be a Hermitian operator, i.e. for two arbitrary functions ψ_1, ψ_2 which have to satisfy certain regularity conditions only, we must have just as in eq. (4.4):

$$\int \psi_1^* H \psi_2 dq = \int \psi_2 [H \psi_1]^* dq.$$

3. As a generalisation of eqs. (3.1'') and (3.4) we assume that

$$\varphi(p_1 \dots p_f; t) = \frac{1}{\sqrt{(2\pi\hbar)^f}} \int \psi(q_1 \dots q_f; t) e^{-\frac{i}{\hbar}(p_1 q_1 + \dots + p_f q_f)} dq \quad (5.5)$$

with the inversion

$$\psi(q_1 \dots q_f; t) = \frac{1}{\sqrt{(2\pi\hbar)^f}} \int \varphi(p_1 \dots p_f; t) e^{+\frac{i}{\hbar}(p_1 q_1 + \dots + p_f q_f)} dp \quad (5.5')$$

so that

$$W(p_1 \dots p_f; t) dp = \varphi \varphi^* dp \quad (5.6)$$

gives the probability of finding, at time t , the momentum of the particles between p_k and $p_k + dp_k$. Here $dq = dq_1 \dots dq_f$ and $dp = dp_1 \dots dp_f$. We have further the completeness relation

$$\int \varphi^* \varphi dp \equiv \int \psi^* \psi dq. \quad (5.7)$$

On integration by parts we obtain, as in the single particle case [eqs. (3.24), (3.24')], the following expectation values:

$$\overline{F(q_1 \dots q_f)} = \int \psi^* F \psi dq = \int \varphi^* \left[F \left(i\hbar \frac{\partial}{\partial p_1}, \dots, i\hbar \frac{\partial}{\partial p_f} \right) \varphi \right] dp. \quad (5.8)$$

$$\overline{F(p_1 \dots p_f)} = \int \varphi^* F \varphi dp = \int \psi^* \left[F \left(\frac{\hbar}{i} \frac{\partial}{\partial q_1}, \dots, \frac{\hbar}{i} \frac{\partial}{\partial q_f} \right) \psi \right] dq. \quad (5.8')$$

Here F denotes a rational integral function of f variables. We shall come back to the consideration of the significance of these relations presently.

As regards the choice of the Hamiltonian operator H , it is assumed that in the case when there is no mutual interaction between the particles, which can, however, be subject to arbitrary external forces, the Hamiltonian operator can be split up into independent summands:

$$H = H^{(1)} + H^{(2)} + \dots + H^{(N)}. \quad (5.9)$$

The operator $H^{(1)}$ alters the function $\psi(x_k^{(1)})$ containing only the co-ordinates of the first particle, leaving the functions involving the co-ordinates of the other particles unaffected. Thus we have

$$H^{(1)} [\psi(q^{(1)}) \psi(q^{(2)} \dots q^{(N)})] = \{H^{(1)} [\psi(q^{(1)})]\} \psi(q^{(2)} \dots q^{(N)}) \dots$$

Similar is the case for the operators $H^{(2)} \dots H^{(N)}$. If, therefore,

$$\psi^{(1)}(q^{(1)}), \dots, \psi^{(N)}(q^{(N)}) \dots$$

are arbitrary solutions of the wave equations

$$-\frac{\hbar}{i} \frac{\partial \psi^{(a)}}{\partial t} = \mathbf{H}^{(a)} \psi^{(a)}, \quad a = 1, 2, \dots, N$$

of the isolated systems, then

$$\psi = \psi^{(1)} \cdot \psi^{(2)} \cdots \psi^{(N)} \quad (5.10)$$

is a solution (of course, not the most general solution) of the equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \mathbf{H} \psi = [\mathbf{H}^{(1)} + \mathbf{H}^{(2)} + \cdots + \mathbf{H}^{(N)}] \psi.$$

An additive decomposition of the Hamiltonian into independent summands corresponds, therefore, to a product decomposition of the wave function into independent factors. This is in agreement with the statement made earlier that for statistically independent particles the probability $W(q_1 \dots q_N; t)$ is split up into a product. Since ψ is uniquely determined for all times by its behaviour at a particular time t_0 , we can say: If the wave function for non-interacting particles can be split up into a product at a given instant of time, then this is true for all times. In other words: if particles which are not mechanically coupled are statistically independent at a particular instant of time t_0 they remain so for all times.

From the previous section we therefore know the Hamiltonian \mathbf{H}_0 for non-interacting particles subject to external forces. It is given by

$$\mathbf{H}_0 = \sum_{a=1}^N \left[-\frac{\hbar^2}{2m^{(a)}} \sum_{k=1}^3 \left(\frac{\partial}{\partial x_k^{(a)}} - \frac{i}{\hbar} \frac{e^{(a)}}{c} \Phi_k^{(a)}(x_i^{(a)}) \right)^2 + V^{(a)}(x_i^{(a)}) \right]. \quad (5.11)$$

If the forces between the particles can be derived from a potential, which depends only on their position co-ordinates and can be written as $V(q_1 \dots q_p)$, it is but proper to set

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \mathbf{H} \psi = \mathbf{H}_0 \psi + V(q_1 \dots q_p) \psi. \quad (5.12)$$

The Coulomb forces between electrically charged particles, the potential of which is given by

$$V = \sum'_{(a, b)} \frac{e_a e_b}{r_{ab}} \quad (5.13)$$

come under this category. [In the sum (5.13) a is different from b and each pair (a, b) is to be taken only once.] The problem of the magnetic interaction between two particles will be dealt with only when we consider relativistic quantum mechanics.

The relations (5.11) to (5.13) for the non-relativistic wave equation of the many-body problem contain, apart from a necessary addition of spin (cf. Sec. 13), the basis for the quantitative treatment of atomic and molecular structures. It is to be stressed that the potentials $\Phi_k^{(a)}$, $V^{(a)}$ and V occurring in them are taken over from the classical theory,³ in particular this is true for the Coulomb potential (5.13) which in turn is a consequence of the Maxwell equations. Thus the present-day wave

³ The “external forces” are to be considered as auxiliary concepts, the employment of which is practical only if the bodies responsible for these forces are not included in the system considered. These forces could be eliminated in principle if the retardation of the forces could be rigorously taken into account in quantum mechanics.

mechanics rests on two basic assumptions: (1) The equation for the matter waves (interpreted only symbolically) which is to be considered as a logical generalisation of classical particle mechanics involving the quantum of action; (2) Maxwell's electrodynamical equations, which, however, likewise require a quantum-mechanical interpretation (see Chapter X).

Except in Chapter X, we shall consider the potentials as given space-time functions. The continuity equation (3.14) and eq. (3.22) for the temporal variation of current can then be taken over directly to our case. For this purpose, we introduce the notation e_k, m_k, Φ_k with $k = 1, 2, \dots, f$, in the place of $e^{[a]}, m^{[a]}, \Phi_k^{[a]}(x^{[a]})$ where $k = 1, 2, 3$ and $a = 1, \dots, N$ so that, e.g., $m_1 = m_2 = m_3 = m^{(1)}$; $m_4 = m_5 = m_6 = m^{(2)}$. We have then, in the f -dimensional co-ordinate space, a current vector i_k with f components ($k = 1, \dots, f$). The physical significance of the current vector is that, e.g., i_1 gives the probability that for given positions of all particles the first particle passes through a unit surface area perpendicular to x_1 in the direction $-x_1$ to $+x_1$ rather than in the reverse direction. The vector \vec{i} in the f -dimensional co-ordinate space is given by

$$i_k = \frac{\hbar}{2m_k i} \left(\psi^* \frac{\partial \psi}{\partial q_k} - \psi \frac{\partial \psi^*}{\partial q_k} \right) - \frac{e_k}{m_k c} \Phi_k \psi^* \psi, \quad (5.14)$$

and it satisfies the continuity equation

$$\frac{\partial (\psi^* \psi)}{\partial t} + \sum_{k=1}^f \frac{\partial i_k}{\partial q_k} = 0. \quad (5.15)$$

The stress-tensor in f -dimensional space is constructed similarly and is given by an expression analogous to (4.20):

$$T_{\chi\lambda} = \frac{\hbar^2}{4m_\lambda} \left[-\psi^* \left(\frac{\partial}{\partial q_\lambda} - \frac{i e_\lambda}{\hbar c} \Phi_\lambda \right) \left(\frac{\partial \psi}{\partial q_\chi} - \frac{i e_\chi}{\hbar c} \Phi_\chi \psi \right) \right. \\ \left. - \psi \left(\frac{\partial}{\partial q_\lambda} + \frac{i e_\lambda}{\hbar c} \Phi_\lambda \right) \left(\frac{\partial \psi^*}{\partial q_\chi} + \frac{i e_\chi}{\hbar c} \Phi_\chi \psi^* \right) \right\} \\ + \left(\frac{\partial \psi}{\partial q_\chi} - \frac{i e_\chi}{\hbar c} \Phi_\chi \psi \right) \left(\frac{\partial \psi^*}{\partial q_\lambda} + \frac{i e_\lambda}{\hbar c} \Phi_\lambda \psi^* \right) \\ + \left. \left(\frac{\partial \psi^*}{\partial q_\chi} + \frac{i e_\chi}{\hbar c} \Phi_\chi \psi^* \right) \left(\frac{\partial \psi}{\partial q_\lambda} - \frac{i e_\lambda}{\hbar c} \Phi_\lambda \psi \right) \right]. \quad (5.16)$$

The symmetry condition $T_{\chi\lambda} = T_{\lambda\chi}$ is valid only if χ and λ refer to the same particle. Similar to (4.19) we then have

$$m \frac{\partial i_\chi}{\partial t} = - \sum_{\lambda} \frac{\partial T_{\chi\lambda}}{\partial q_\lambda} + \left\{ - \frac{\partial (V + \sum_a V^{(a)})}{\partial q_\chi} - \frac{e_\chi}{c} \frac{\partial \Phi_\chi}{\partial t} \right\} \psi^* \psi \\ + \frac{e_\chi}{c} \sum_{\lambda} \left(\frac{\partial \Phi_\lambda}{\partial q_\chi} - \frac{\partial \Phi_\chi}{\partial q_\lambda} \right) i_\lambda. \quad (5.17)$$

According to our assumption about Φ_k , only three terms in the last sum (which refer to the same particle) are different from zero. Further we have, analogous to (4.16'') and to (4.23)

$$\frac{d\overline{q_k}}{dt} = \int i_k dq = \frac{1}{m_k} \left(\overline{p_k} - \frac{e_k}{c} \overline{\Phi_k} \right), \quad (5.18)$$

$$m \frac{d^2 \overline{q_k}}{dt^2} = \overline{K_k}, \quad (5.19)$$

where \bar{K}_k is defined as in (4.22).

The last mentioned relations – the equations of motion – can be quite generally derived from the wave equation using the operator calculus. We start from the relations (5.5) and (5.5') from which the relations (5.8) and (5.8') follow, if F denotes a rational integral function. This leads to an association of operators to the momenta and co-ordinates which act in the following way:

$$\mathbf{p}_k \psi(q_1 \dots q_l) = \frac{\hbar}{i} \frac{\partial}{\partial q_k} \psi; \quad q_k \psi(q_1 \dots q_l) = q_k \psi; \quad (5.20)$$

$$\mathbf{p}_k \varphi(p_1 \dots p_l) = \mathbf{p}_k \varphi(p_1 \dots p_l); \quad q_k \varphi(p_1 \dots p_l) = -\frac{\hbar}{i} \frac{\partial}{\partial p_k} \varphi. \quad (5.20')$$

From these we obtain the fundamental *commutation relations* (hereafter abbreviated as C.R.):

$$\left. \begin{aligned} \mathbf{p}_k \mathbf{q}_l - \mathbf{q}_k \mathbf{p}_l &= \delta_{lk} \frac{\hbar}{i}, & \delta_{lk} &= \begin{cases} 1 & \text{for } l = k \\ 0 & \text{for } l \neq k \end{cases} \\ \mathbf{p}_k \mathbf{p}_l - \mathbf{p}_l \mathbf{p}_k &= 0, \\ \mathbf{q}_k \mathbf{q}_l - \mathbf{q}_l \mathbf{q}_k &= 0. \end{aligned} \right\} \quad (5.21)$$

For example,

$$\mathbf{p}_k \mathbf{q}_k \psi = \frac{\hbar}{i} \frac{\partial}{\partial q_k} (q_k \psi); \quad \mathbf{q}_k \mathbf{p}_k \psi = q_k \frac{\hbar}{i} \frac{\partial}{\partial q_k} \psi,$$

so that, indeed,

$$(\mathbf{p}_k \mathbf{q}_k - \mathbf{q}_k \mathbf{p}_k) \psi = \frac{\hbar}{i} \left(\frac{\partial}{\partial q_k} (q_k \psi) - q_k \frac{\partial \psi}{\partial q_k} \right) = \frac{\hbar}{i} \psi.$$

We would have obtained the same result had we used the function $\varphi(p_1 \dots p_l)$ for verification of the C.R. Similarly we verify the other C.R. of eq. (5.21). This form of C.R. is just another expression for the relations (5.5) and (5.5') between $\varphi(p)$ and $\psi(q)$.

It is to be stressed that p_k and q_k are Hermitian (linear) operators. These are defined by relations resembling (4.4'):

$$\int \psi_1^* (\mathbf{H} \psi_2) dq = \int \psi_2 (\mathbf{H} \psi_1)^* dq. \quad (5.22)$$

This is valid for arbitrary functions ψ_1 and ψ_2 and can be easily verified for the operators (5.20). Further, on applying (5.22) twice on two Hermitian operators, we get

$$\begin{aligned} \int (\mathbf{H}_1 \psi_1)^* (\mathbf{H}_2 \psi_2) dq &= \int \psi_2 (\mathbf{H}_2 [\mathbf{H}_1 \psi_1])^* dq, \\ \int (\mathbf{H}_2 \psi_2)^* (\mathbf{H}_1 \psi_1) dq &= \int \psi_1 (\mathbf{H}_1 [\mathbf{H}_2 \psi_2])^* dq. \end{aligned}$$

Hence

$$\int \psi_2 (\mathbf{H}_2 [\mathbf{H}_1 \psi_1])^* dq = \int \psi_1^* (\mathbf{H}_1 [\mathbf{H}_2 \psi_2]) dq. \quad (5.23)$$

From this we have: if \mathbf{H}_1 and \mathbf{H}_2 are linear Hermitian operators, the same holds true for

$$\mathbf{F} = \mathbf{H}_1 \mathbf{H}_2 + \mathbf{H}_2 \mathbf{H}_1 \quad (5.24)$$

and

$$\mathbf{G} = i(\mathbf{H}_1 \mathbf{H}_2 - \mathbf{H}_2 \mathbf{H}_1). \quad (5.24')$$

If, in particular, \mathbf{H}_1 and \mathbf{H}_2 commute, then $\mathbf{H}_1 \mathbf{H}_2$ is also Hermitian and every rational

integral function of H , is again Hermitian. If A, B are two linear operators, we often write for brevity

$$[A, B] = i(AB - BA). \quad (5.25)$$

Then we have

$$[A_1 A_2, A_3] = A_1 [A_2 A_3] + [A_1 A_3] A_2, \quad (5.26)$$

$$[[A_1, A_2] A_3] + [[A_2, A_1] A_3] + [[A_2, A_3] A_1] = 0. \quad (5.27)$$

Let F be an arbitrary linear Hermitian operator, which does not depend on the time explicitly and H , the Hamiltonian operator. We shall calculate the variation with time of the mean value ("the expectation value"):

$$\bar{F} = \int \psi^* (F\psi) dq. \quad (5.28)$$

We obtain

$$\begin{aligned} \hbar \frac{d\bar{F}}{dt} &= \hbar \int \frac{\partial \psi^*}{\partial t} (F\psi) dq + \hbar \int \psi^* \left(F \frac{\partial \psi}{\partial t} \right) dq \\ &= i \int (H\psi)^* (F\psi) dq - i \int \psi^* (F[H\psi]) dq \\ &= i \int \psi^* [(HF)\psi] dq - i \int \psi^* [(FH)\psi] dq, \end{aligned}$$

i.e.

$$\hbar \frac{d\bar{F}}{dt} = \int \psi^* ([H, F]\psi) dq = \overline{[H, F]} = i(\overline{HF} - \overline{FH}). \quad (5.29)$$

Now, for every function $F(p_1 \dots p_f)$ of the momenta alone, we have

$$F p_k - p_k F = 0; \quad F q_k - q_k F = \frac{\hbar}{i} \frac{\partial F}{\partial p_k}. \quad (5.30)$$

The latter formula is correct for $F = p_i$ and for $F = q_i$; it is also correct for $F_1 + F_2$ and $F_1 \cdot F_2$, if it holds for F_1 and F_2 separately, as we can verify from eq. (5.26). From this follows the generalisation of eq. (5.30) for every rational integral function F of p . Further according to the definition $p_k = \frac{\hbar}{i} \frac{\partial}{\partial q_k}$, we have for every function $G(q_1 \dots q_f)$ of co-ordinates alone

$$p_k G - G p_k = \frac{\hbar}{i} \frac{\partial G}{\partial q_k}; \quad q_k G - G q_k = 0. \quad (5.31)$$

From eqs. (5.30) and (5.31) taken together, it follows, first for every function

$$H(p, q) = F(p_1 \dots p_f) + G(q_1 \dots q_f), \quad (5.32)$$

where F is a rational integral function and G is arbitrary, that is

$$H\psi(q) = \left[F\left(\frac{\hbar}{i} \frac{\partial}{\partial q_1}, \dots, \frac{\hbar}{i} \frac{\partial}{\partial q_f}\right) + G(q_1 \dots q_f) \right] \psi,$$

$$H p_k - p_k H = -\frac{\hbar}{i} \frac{\partial H}{\partial q_k}; \quad H q_k - q_k H = \frac{\hbar}{i} \frac{\partial H}{\partial p_k}. \quad (5.33)$$

Finally from the definition of p_k and q_k , it is easy to prove the above formula for a function of the form

$$H = F(p_1 \dots p_f) + \sum_k [A_k(q) p_k + p_k A_k(q)] + G(q_1 \dots q_f). \quad (5.32')$$

The Hamiltonian in cartesian co-ordinates, which we have used till now, is of this form. We have to pay attention to the symmetrisation of the order of the factors A_k and p_k which is necessary, according to eq. (5.24), in order that H is Hermitian.

On substituting $F = p_k$ and $F = q_k$ successively in eq. (5.29), we obtain, taking into consideration eq. (5.33), the mean values

$$\frac{d\bar{p}_k}{dt} = - \overline{\left(\frac{\partial H}{\partial q_k} \right)}; \quad \frac{d\bar{q}_k}{dt} = + \overline{\left(\frac{\partial H}{\partial p_k} \right)} \quad (5.34)$$

taken over the wave-packet constructed from an arbitrary solution of the wave equation. Here, in generalisation of the above, the mean value of an expression of the form

$$A_k(q) \bar{p}_k + \bar{p}_k A_k(q)$$

is meant to be the real value

$$\int \psi^* \left[A_k(q) \frac{\hbar}{i} \frac{\partial \psi}{\partial q_k} + \frac{\hbar}{i} \frac{\partial}{\partial q_k} (A_k(q) \psi) \right] dq$$

This definition has proved appropriate in many contexts. Since $[H, H] = 0$, it follows from (5.29) with $F = H$ that if H does not contain time explicitly, we have

$$\frac{d\bar{H}}{dt} = 0; \quad \bar{H} = \text{const.} \quad (5.35)$$

We recognise in this expression the law of conservation of energy, since \bar{H} can be interpreted as the mean value of the energy taken over the wave-packet. Similarly for the total momentum we have

$$\bar{P} = \sum_{(k)} \bar{p}_k; \quad \frac{d\bar{P}}{dt} = - \overline{\left(\sum_k \frac{\partial}{\partial q_k} \right) H},$$

which vanishes if H depends explicitly on the difference of the co-ordinates, $q_k - q_i$, only. Further in the absence of a magnetic field, we have for the angular momentum

$$J_{ik} = \sum_{a=1}^N (q_i^{(a)} p_k^{(a)} - q_k^{(a)} p_i^{(a)}), \quad (J_{ik} = -J_{ki}; i, k = 1, 2, 3) \quad (5.36)$$

and

$$\frac{d\bar{J}_{ik}}{dt} = - \sum_{a=1}^N \overline{\left[q_i^{(a)} \frac{\partial V}{\partial q_k^{(a)}} - q_k^{(a)} \frac{\partial V}{\partial q_i^{(a)}} \right]}, \quad (5.37)$$

[(a) = particle index going from 1 to N] where, as in classical mechanics, the right-hand side of eq. (5.37) vanishes, if the potential energy of the system is invariant under a rigid rotation of the whole system in space. When a magnetic field is present, it follows immediately that

$$\frac{d\bar{q}_k}{dt} = \frac{1}{m} \left(\bar{p}_k - \frac{e}{c} \bar{\Phi}_k \right) = \int i_k dq, \quad (5.38)$$

if i_k is defined by (5.14). Further, with $\mathcal{H}_{\lambda} = \frac{\partial \Phi_{\lambda}}{\partial q_k} - \frac{\partial \Phi_{\lambda}}{\partial q_k}$, from our definition of the mean value, follows

$$\int \mathcal{H}_{\lambda} i_k dV = \frac{1}{2} \overline{(\mathcal{H}_{\lambda} \dot{q}_{\lambda} + \dot{q}_{\lambda} \mathcal{H}_{\lambda})} = \frac{1}{2m} \overline{(\mathcal{H}_{\lambda} p_{\lambda} + p_{\lambda} \mathcal{H}_{\lambda})} - \frac{e}{mc} \mathcal{H}_{\lambda} \Phi_{\lambda}.$$

Hence

$$m \frac{d^2 q_x}{dt^2} = -\frac{\partial (V + \sum_a V^{(a)})}{\partial q_x} - \frac{e_x}{c} \frac{\partial \Phi_x}{\partial t} + \frac{e_x}{c} \frac{1}{2} \sum_k (\mathcal{H}_{x\lambda} \dot{q}_\lambda + \dot{q} \mathcal{H}_{x\lambda}) = \overline{K}_x$$

and with

$$\left. \begin{aligned} J_{ik} &= \sum_{a=1}^N m^{(a)} (q_i^{(a)} \dot{q}_k^{(a)} - q_k^{(a)} \dot{q}_i^{(a)}) \\ &= \sum_{a=1}^N \left[(q_i^{(a)} p_k^{(a)} - p_k^{(a)} q_i^{(a)}) - \frac{e^{(a)}}{c} (q_i^{(a)} \Phi_k^{(a)} - q_k^{(a)} \Phi_i^{(a)}) \right], \end{aligned} \right\} \quad (5.36')$$

also follow

$$\overline{J}_{ik} = \sum_{a=1}^N m^{(a)} \int [q_i^{(a)} \dot{q}_k^{(a)} - q_k^{(a)} \dot{q}_i^{(a)}] dq, \quad (5.36'')$$

and

$$\frac{d \overline{J}_{ik}}{dt} = \frac{1}{2} \sum_{(a)} [(q_i^{(a)} K_k^{(a)} - q_k^{(a)} K_i^{(a)}) + (K_k^{(a)} q_i^{(a)} - K_i^{(a)} q_k^{(a)})], \quad (5.37')$$

where by $K_k^{(a)}$ we understand the operator corresponding to the component of the force concerned. The last equation follows also directly from eq. (5.17). The right-hand side of (5.37') vanishes if the system possesses rotational symmetry about an axis perpendicular to the $(x_i x_k)$ -plane (cf. Sec. 13).

We have yet to consider the case where we use co-ordinates other than cartesian. Since the classical Hamiltonian there is quadratic in p , but with arbitrary coefficients depending on q , there is, in general, an ambiguity regarding the order of the factors $f(q)$ and p_k . This order cannot be determined except by changing back to cartesian co-ordinates.⁴ On the other hand, it is possible to give a rational prescription for the construction of partial differential coefficients with respect to p_k and q_k of such general expressions.⁵ That is, we can stipulate that in general for a product of two functions, $F_1 \cdot F_2$, the order of the factors should be

$$\frac{\partial}{\partial X} (F_1 F_2) = \frac{\partial F_1}{\partial X} F_2 + F_1 \frac{\partial F_2}{\partial X}, \quad (5.39)$$

from which, by induction, it follows for any finite number of factors

$$\frac{\partial}{\partial X} (F_1 \dots F_N) = \frac{\partial F_1}{\partial X} F_2 \dots F_N + F_1 \frac{\partial F_2}{\partial X} F_3 \dots F_N + \dots + F_1 \dots F_{n-1} \frac{\partial F_n}{\partial X}, \quad (5.39')$$

where X stands for any of the variables $p_1 \dots q_r$. In this case, eqs. (5.33) are, in general, correct, if H is a rational integral function of p and an arbitrary function of q . Again on account of eq. (5.29), eq. (5.34) is then a consequence of the wave equation.

We can now formulate the wave equation in any arbitrary curvilinear co-ordinates. Let the line element be

$$ds^2 = g_{\alpha\lambda} dq_\alpha dq_\lambda$$

(a summation over all repeated indices is always understood in the following equations), where $g_{\alpha\lambda} = g_{\lambda\alpha}$ are arbitrary functions of q_λ and the scale factor has been absorbed in $g^{\alpha\lambda}$. The $g^{\alpha\lambda}$ are matrices reciprocal to $g_{\alpha\lambda}$ and $D = \sqrt{|g|}$ is the square root of the determinant $|g| = |g_{\alpha\lambda}|$. In these co-ordinates the wave equation corresponding to eqs. (5.11) and (5.12) then reads

⁴ Cf. B. Podolsky, Phys. Rev. 32, 812 (1928).

⁵ M. Born, P. Jordan and W. Heisenberg, Z. Physik 35, 557 (1926).

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \frac{1}{2} \frac{1}{D} \left(\frac{\hbar}{i} \frac{\partial}{\partial q_x} + A_x \right) D g^{x\lambda} \left(\frac{\hbar}{i} \frac{\partial}{\partial q_\lambda} + A_\lambda \right) \psi + V \psi, \quad (5.40)$$

where A_x are the vector potentials multiplied by $-\frac{e_x}{c}$. With

$$\rho = D \psi \psi^*,$$

$$i^* = D g^{x\lambda} \left[\psi^* \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial q_\lambda} + A_\lambda \psi \right) + \psi \left(-\frac{\hbar}{i} \frac{\partial \psi^*}{\partial q_\lambda} + A_\lambda \psi^* \right) \right]$$

we have the continuity equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial i^*}{\partial q_x} = 0.$$

On account of the occurrence of the factor D in the density function, an operator F is called Hermitian, if

$$\int D \psi^* (F \psi) dq = \int D (F \psi)^* \psi dq.$$

If the momentum operator p_x is to be Hermitian in this sense and if further it satisfies the C.R.

$$p_x q_x - q_x p_x = \frac{\hbar}{i}$$

we must have

$$p_x \psi = \frac{\hbar}{i} \sqrt{D} \frac{1}{\sqrt{D}} \frac{\partial \sqrt{D} \psi}{\partial q_x}.$$

The relationship of these operators with the wave equation and the current can be easily derived.

In the special case of spherical polar co-ordinates we have

$$ds^2 = m(d\tau^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2),$$

so that

$$D = r^2 \sin \theta, \quad g^{rr} = \frac{1}{m}, \quad g^{\theta\theta} = \frac{1}{m} \frac{1}{r^2}, \quad g^{\varphi\varphi} = \frac{1}{m} \frac{1}{r^2 \sin^2 \theta}.$$

For later applications it is to be remembered that in this case

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{df}{dr} \right) = \frac{1}{r} \frac{d^2 f}{dr^2} (rf).$$

Hence we can, on taking account of eq. (5.40) and using the definitions

$$p_r \psi = \frac{\hbar}{i} \frac{1}{r} \frac{\partial}{\partial r} (r \psi) \quad \text{and} \quad \mathbf{P}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) \quad (5.41)$$

write the Hamiltonian simply as

$$\mathbf{H} = \frac{1}{2m} \left(\mathbf{p}_r^2 + \frac{\mathbf{P}^2}{r^2} \right) + V. \quad (5.42)$$

This way of writing is often used in the older works on quantum mechanics.

CHAPTER III

Stationary States and the Eigenvalue Problem

6. Stationary States as Eigenvalue Problem

Among the solutions of the general wave equation (4.3)

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \mathbf{H} \left(\frac{\hbar}{i} \frac{\partial}{\partial q}, q \right) \psi , \quad (6.1)$$

those for which the probability density $\psi^* \psi$ as well as the current density \vec{j} are constant in time, are of special interest. From now on we assume that the scalar and vector potentials V and Φ_k , appearing in \mathbf{H} , do not depend on time explicitly. In this case, we say that the solutions correspond to the *stationary states* of the system. In order that $\psi^* \psi$ as well as $\psi^* \frac{\partial \psi}{\partial q_k} - \psi \frac{\partial \psi^*}{\partial q_k}$ be independent of time the wave function ψ must necessarily have the form

$$\psi = u(q) e^{-i/Et} , \quad (6.1a)$$

where $u(q)$ is independent of t and E is independent of q . On substitution of (6.1a) in eq. (6.1), it follows that

$$\hbar \frac{df}{dt} u(q) = \mathbf{H}[u(q)]$$

and this is possible only if df/dt is independent of t . We can, therefore, write (6.1a) as

$$\psi_E = u(q) e^{-\frac{i}{\hbar} Et} . \quad (6.2)$$

Substituting (6.2) in (6.1) we obtain

$$\mathbf{H} \left(\frac{\hbar}{i} \frac{\partial}{\partial q}, q \right) u = E u . \quad (6.3)$$

This is a homogeneous, linear differential equation containing a parameter. As is well known, such differential equations do not always have regular solutions for all values of E and we, therefore, have an *eigenvalue problem*.* In order to study the regularity conditions of the problem,¹ we start from the equation, valid for any two arbitrary

*See, e.g., R. Courant and D. Hilbert: Methods of Mathematical Physics, Vol. I, Interscience Publishers, Inc., New York, (1975), Chapter V.

¹Cf. J.v. Neumann, Göttinger Nach. (1927), 1. This question was again discussed by G. Jaffé, Z. Physik 66, 770 (1930). It appears to us, however, that the most general answer to the question is given in the above paper of von Neumann.

functions u, v :

$$\int v^*(H u) dq = \int u(H v)^* dq. \quad (6.4)$$

Only those functions, which do not disturb the "Hermiticity" of H by having singular points, are permitted to be solutions of eq. (6.3). In particular, relation (6.4) should hold, if one of the two functions is regular. We can make still more general assumptions regarding the domain of values of q and the range of the function u . [In general dq denotes the volume element multiplied by a suitable density function $p(q)$: $dq = p(q)dq_1 \dots dq_n$.] For example, for the rotation of rigid bodies we have to deal with angles which can vary only from 0 to 2π or from 0 to π . Alternatively, we may have to consider functions in a bounded interval, say $[-1, +1]$, which take the same value at -1 and $+1$ or to consider functions which are defined in the half-space $(0, \infty)$ and have to vanish at $q = 0$. The relation (6.4) must always be satisfied in the region of validity of the function which is subject to certain boundary and regularity conditions. For this purpose, it is not necessary that u and v should be everywhere regular. But on account of the interpretation of $\psi\psi^* = uu^*$ as probability densities, it is natural to demand that

$$\int u u^* dq \leq K \quad (K < \infty), \quad (6.5)$$

where the integration is over the whole domain of q . In many cases this requirement gives rise to *discrete* eigenvalues E . The requirement (6.5) has to be somewhat weakened for continuous eigenvalues, as can be seen for the case of free particles. A solution of the equation for this case,

$$-\frac{\hbar^2}{2m} \Delta u = E u$$

is the plane wave

$$u_{p_1 \dots p_f} = e^{\frac{i}{\hbar}(p_1 q_1 + \dots + p_f q_f)},$$

where p_1, \dots, p_f are integration constants. Evidently the condition (6.5) is not satisfied for these plane waves. In fact, the plane waves represent physically a singular limiting case, in which the probability of finding a particle outside a finite volume is infinitely larger than the probability of finding it inside. Only the ratio of the probabilities for finding a particle in two different finite volumes is physically significant.² As already shown in Sec. 2 (Chapter I) the singularity mentioned above disappears, when we consider a wave-packet. If we construct a wave-packet by integration over a finite but arbitrarily small region of the p -space, i.e.

$$\bar{u}_{p'_k, p''_k} = \int_{p_k}^{p''_k} dp_1 \dots dp_f u_{p_1 \dots p_f},$$

then the integral of $\bar{u} \bar{u}^*$ over the whole of q -space exists. In the case of eigenfunctions depending continuously on arbitrary parameters $\lambda_1, \lambda_2, \dots$ (in short, denoted as λ) we write

$$\bar{u}_{\lambda', \lambda''} = \int_{\lambda'}^{\lambda''} u_{\lambda}(q) d\lambda, \quad (6.6)$$

² Cf. for this P.A.M. Dirac, The Principles of Quantum Mechanics, Fourth edn., Oxford University Press, (1959), p. 185.

where $d\lambda = d\lambda_1 \dots d\lambda_n$. Then we have to require instead of (6.5) that

$$\int \bar{u}_{\lambda, \lambda'}^*(q) \bar{u}_{\lambda, \lambda'}(q) dq \leq K', \quad (K' < \infty). \quad (6.7)$$

Only for these elementary wave-packets $\bar{u}_{\lambda, \lambda'}$ need then the Hermiticity relation (6.4) hold for the whole of q -space.

We obtain a unified formulation of the discrete and continuous cases by setting

$$\bar{u}_\lambda = \int_{\lambda_0}^{\lambda} u_\lambda(q) dq + \sum_{\lambda_p=\lambda_1}^{\lambda_n} u_p(q), \quad (6.6')$$

where the integration is over the continuous eigenvalues lying within the interval $\lambda_0 \leq \lambda' < \lambda$ and the summation is over the discrete eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Thus \bar{u}_λ can change with λ discontinuously. The Stieltjes integral

$$\int_{\lambda_0}^{\lambda} f(\lambda) d\bar{u}_\lambda = \int_{\lambda_0}^{\lambda} f(\lambda) u_\lambda(q) dq + \sum_{\lambda_p=\lambda_1}^{\lambda_n} u_p f(\lambda_p)$$

exists for any continuous function $f(\lambda)$ of this type. We can now (quite independently from the existence of the u_λ) characterise the function \bar{u}_λ , which may possibly vary with λ discontinuously, by the equation

$$(H \bar{u}_{\lambda'}) - (H \bar{u}_\lambda) = \int_{\lambda'}^{\lambda''} [H d\bar{u}_\lambda] = \int_{\lambda'}^{\lambda''} E(\lambda) d\bar{u}_\lambda. \quad (6.3')$$

This can be considered as a generalisation of (6.3) and, at the points of discontinuity $\lambda_1, \lambda_2, \dots$ of \bar{u}_λ , the values of E appear in it as discrete eigenvalues E_1, E_2, \dots

In order to investigate the singularities of u and v , which are compatible with (6.4), it is useful to start from the equation

$$\int [v^* H u - u (H v)^*] dq = \oint i_N(v^*, u) d\vec{l} \quad (6.8)$$

which follows from the continuity equation

$$v^* H u - u (H v)^* = \text{Div } \vec{i}(v^*, u)$$

by integrating over a finite volume of the q -space. $\vec{i}(v^*, u)$ is the usual current density from which we obtain $\vec{i}(v^*, u)$ by the formal substitutions $v^* \rightarrow v^*$, $v \rightarrow u$; the notation 'Div' above refers to divergence in an f -dimensional space (possibly represented in curvilinear co-ordinates). The surface integral on the right-hand side of eq. (6.8) in which i_N denotes the component of \vec{i} in the direction of the outward normal, extends over the bounding surface enclosing the volume V . If, in particular, the left-hand side of eq. (6.8) represents a one-dimensional integral, then the right-hand side of eq. (6.8) reduces to the difference of the values of \vec{i} at the limits of integration q_1, q_2, \dots . At the places where the potentials occurring in H are singular, we must require that the surface integrals appearing in eq. (6.8), when we enclose the singularities, can be made arbitrarily small.

From the above we infer easily the following:

(a) Let a one-dimensional problem with the Hamiltonian

$$\frac{1}{2m} \left(\frac{\hbar}{i} \frac{d}{dx} - \frac{e}{c} \Phi(x) \right)^2 + V(x)$$

be given with either $\Phi(x)$ or $V(x)$ or both being discontinuous at a point x_0 . Since

$$i(v^*, u) = \frac{1}{2m} \left[v^* \left(\frac{\hbar}{i} \frac{du}{dx} - \frac{e}{c} \Phi(x) u \right) - u \left(\frac{\hbar}{i} \frac{dv^*}{dx} + \frac{e}{c} \Phi(x) v^* \right) \right],$$

we require

$$\frac{du}{dx} - \frac{i\hbar e}{c} \Phi(x) u \text{ continuous, } u \text{ continuous at } x = x_0. \quad (6.9)$$

These conditions must hold for all the functions to be considered here and hence also for $v(x)$.

(b) Let the potential V in ordinary space in which a particle moves be singular. The integral $\oint_{i_N} df$ taken over a small sphere around the origin is

$$r^2 \int [v^* \left(\frac{\hbar}{i} \frac{\partial u}{\partial r} - \frac{e}{c} \Phi, u \right) - u \left(\frac{\hbar}{i} \frac{\partial v^*}{\partial r} + \frac{e}{c} \Phi, v^* \right)] d\Omega.$$

If all the functions v and u are of the form \bar{v}/r^α , \bar{u}/r^α respectively where \bar{v} , \bar{u} are regular at the origin (and perhaps vanish there, but this is not necessary) then we see that we must have $2 - 2\alpha > 0$ and hence $\alpha < 1$, in order that this integral vanishes for all regular v , u in the limit $r \rightarrow 0$, i.e. v , u must tend to infinity more slowly than $1/r$. An eigenfunction for which $\lim (ru) = A \neq 0$, is not admissible (though for such a function $\int_0^\infty u^* u r^2 dr$ exists).

We have thus a well-defined eigenvalue problem and hence a natural and definite method for determining the possible discrete or continuous eigenvalues of the system. This method originated with Schrödinger,³ who was able to show that for an electron with the charge $-e$ moving in the field of a fixed nucleus of charge $+Ze$ (hydrogen-like atom), for which the potential energy is $V = -(Ze^2/r)$, the negative, discrete energy eigenvalues $-Rh/1^2, -Rh/2^2, \dots, -Rh/n^2, \dots$ and the adjacent, continuous, positive energy eigenvalues ($0 \leq E < +\infty$) follow from his wave equation [cf. (5.12), (5.13)]

$$-\frac{\hbar^2}{2m} \Delta u + \left(E + \frac{Z e^2}{r} \right) u = 0.$$

In the above

$$R = \frac{m e^4}{4\pi \hbar^3} Z^2$$

is Rydberg's constant multiplied by the square of the charge of the nucleus. The mixing of the discrete and continuous eigenvalues here is essentially due to the behaviour of the Coulomb potential V at large distances, namely, the slow increase of the negative values towards zero.* On the other hand the singularity of V at the origin is quite inessential, since the point nucleus can as well be replaced by a small charged sphere.

The simplest example of a system with only discrete eigenvalues is the harmonic oscillator which possesses, in the one-dimensional case, the potential energy

$$V = \frac{m}{2} \omega^2 x^2,$$

and in the three-dimensional case, if there is isotropy, the potential energy

$$V = \frac{m}{2} \omega^2 r^2$$

³ See references in footnote 4, Sec. 1.

* von Neumann and Wigner and more recently Weidmann have constructed explicit examples of potentials which possess positive energy eigenvalues corresponding to bound states: J. von Neumann and E.P. Wigner, Z. Physik **30** (1929), pp. 465-67; J. Weidmann, Math. Zeit. **98** (1967), pp. 268-302.

where ω is the angular frequency of the oscillator. The energy eigenvalues are

$$E_n = (n + \frac{1}{2}) \hbar \omega \quad n = 0, 1, 2, 3, \dots$$

for a linear oscillator, and

$$E_n = (n + \frac{3}{2}) \hbar \omega \quad n = 0, 1, 2, 3, \dots$$

for a three-dimensional isotropic oscillator.* In the cases considered here (just as in the case of a free particle) the sequence of energy values is bounded from below, i.e. by $\frac{1}{2}(\hbar\omega)$ for the linear and $3/2(\hbar\omega)$ for the isotropic oscillator. That this need not always be the case is shown by the example**

$$V = -F x,$$

which corresponds to a constant force in the $+x$ direction.

As an example of a calculation in curvilinear co-ordinates we shall go through the discussion of the special case of polar co-ordinates which is important for the problem of the hydrogen atom, and in fact for any central field problem, i.e. motion of a particle in a field, the potential energy $V(r)$ of which depends only on the distance r of the particle from a fixed centre [cf. eqs. (5.41) and (5.42) of Sec. 5]. The operator Δ assumes, in polar co-ordinates (r, θ, ϕ) , the form

$$\Delta u = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r u) + \frac{1}{r^2} \Omega u,$$

where

$$\Omega = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

is connected in a simple way with the operator for the angular momentum of the particle. The components of the latter are given by

$$\mathbf{P}_1 = \frac{\hbar}{i} \left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2} \right)^{\dagger},$$

the others being obtained by cyclic permutation. The square of the angular momentum vector

$$\mathbf{P}^2 = \mathbf{P}_1^2 + \mathbf{P}_2^2 + \mathbf{P}_3^2$$

is then simply

$$\mathbf{P}^2 = -\hbar^2 \Omega.$$

For a central field the solution of the wave equation (6.3) is separable in the form

$$u = f(r) Y(\theta, \phi),$$

where

$$\Omega Y = \lambda Y \tag{6.10}$$

λ being a pure number. The radial part $f(r)$ satisfies the equation

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r} \frac{d^2}{dr^2} (r f) + \frac{\lambda}{r^2} f \right] + V(r) f = E f.$$

*Cf. S. Flügge, Practical Quantum Mechanics, Vol. I, Springer Verlag, Berlin, (1971), Problem 65, p. 168.

**Cf. L.D. Landau and E.M. Lifshitz, Quantum Mechanics, Pergamon Press, London, (1958), p. 70.

[†]In spherical polar co-ordinates

$$\mathbf{P}_1 = i\hbar (\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi}),$$

$$\mathbf{P}_2 = i\hbar (-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi}),$$

It is now important to find the solutions of eq. (6.10) and the eigenvalues of Ω . The answer is well known.⁴ Only for

$$\lambda = -l(l+1), \quad (6.10')$$

where l is a non-negative integer ($l \geq 0$) do solutions of (6.10), which are free from singularities, exist and in fact these solutions are the spherical harmonics of order l . There are $2l+1$ of these eigenfunctions which are linearly independent. The eigenvalues of P^2 are, therefore $\hbar^2 l(l+1)$. If we choose the function $Y_l(\theta, \varphi)$ such that

$$Y_{l,m}(\theta, \varphi) = Y_{l,m}(\theta) e^{im\varphi},$$

then the operator

$$P_3 = \frac{\hbar}{i} \frac{\partial}{\partial \varphi}$$

has $Y_{l,m}(\theta, \varphi)$ as the eigenfunctions corresponding to the eigenvalues $\hbar m$. The integer m runs from $-l$ to $+l$. The functions $Y_{l,m}$ are orthogonal, as can be seen readily, for the same l and different values of m ; for different values of l , however, the orthogonality follows from the differential equation (6.10) read with eq. (6.10'). For two arbitrary functions Y_1 and Y_2 , we have

$$\begin{aligned} \sin \theta (Y_1 \Omega Y_2 - Y_2 \Omega Y_1) &= \frac{\partial}{\partial \theta} \left[\sin \theta \left(Y_1 \frac{\partial Y_2}{\partial \theta} - Y_2 \frac{\partial Y_1}{\partial \theta} \right) \right] \\ &\quad + \frac{\partial}{\partial \varphi} \left[\frac{1}{\sin \theta} \left(Y_2 \frac{\partial Y_1}{\partial \varphi} - Y_1 \frac{\partial Y_2}{\partial \varphi} \right) \right]. \end{aligned}$$

On putting $Y_1 = Y_{lm}$ and $Y_2 = Y_{l'm'}$ we obtain from eq. (6.10)

$$\int Y_{l'm'}^* Y_{lm} \sin \theta d\theta d\varphi = 0 \quad \text{for } l \neq l'.$$

In this connection it is of interest to discuss the possibility of admitting many-valued functions as solutions of (6.10). One could be in doubt whether the requirement of single-valuedness of u is necessary at all, since only the density $\psi^* \psi$, where $\psi = \sum c_n e^{-i/\hbar E t} u_n$, has a direct physical significance. If all the eigenfunctions u_n of the system considered are multiplied by the same factor with modulus 1, then $\psi^* \psi$ remains single-valued, when taken around a closed path.

A general criterion for the admissibility of eigenfunctions, which does not from the beginning assume their single-valuedness, has been given by W. Pauli.^{5*} This criterion states that repeated action of the operators corresponding to physical quantities on the eigenfunctions should not lead outside of the domain of square integrable eigenfunctions. In the example of the spherical harmonic functions considered here, the relevant operators are the components P_1, P_2, P_3 of angular momentum P . The above criterion leads to the exclusion of all spherical functions which are not single-valued, in particular those with half-integral index. The criterion for the admissibility of eigenfunctions here as well as in other examples signifies that these functions must form the basis for a representation of the transformation group (see Sec. 13), which is associated with the eigenvalue problem. In the example considered now, it is the rotation group in three-dimensional space, which is the

⁴ R. Courant and D. Hilbert, loc. cit., p. 317.

⁵ W. Pauli, Helv. Phys. Acta 12, 147 (1939); cf. also M. Fierz, Helv. Phys. Acta 17, 27 (1943).

*Cf. also E. Merzbacher, Am. J. Phys. 30, 237 (1962).

transformation group, for which, as is well known, only the spherical functions with integral index form a (finite-dimensional) representation.

An example for which single-valued eigenfunctions are not allowed by the above criterion is the spherical top.⁶ Here the wave functions depend on the three Eulerian angles θ , φ , ψ . If we form the two complex quantities

$$\alpha = \cos \frac{\theta}{2} e^{\frac{i}{2}(\varphi+\psi)}, \quad \beta = i \sin \frac{\theta}{2} e^{\frac{i}{2}(\varphi-\psi)},$$

defined through eq. (13.26), Sec. 13, the splitting of which into real and imaginary parts yields four real quantities with $|\alpha|^2 + |\beta|^2 = 1$, then it is enough if we require the eigenfunctions to be single-valued on this hyper-spherical surface of the four-dimensional space. This includes the possibility of admitting eigenfunctions which are double-valued in the Euler angles. The eigenfunctions are in fact the harmonic polynomials on this three-dimensional hyper-surface. Group-theoretically these give rise to a direct product of two representations of the rotation group in three-dimensional space, the rotations being respectively in the body-fixed and in the space-fixed co-ordinate systems. The representations with half-integral quantum numbers (Sec. 13) are included here.*

The eigenfunctions have an important property, which follows directly from the basic eq. (6.4) for the operator H . This property is obtained if we substitute for v and u , two solutions u_n , u_m belonging to different energy eigenvalues E_n , E_m respectively. Here we restrict ourselves to the case of discrete eigenvalues in order to apply eq. (6.4) directly to the eigensolutions. Then it follows that⁷

$$E_n \int u_m^* u_n dq = E_m \int u_n u_m^* dq.$$

Therefore

$$\int u_m^* u_n dq = 0 \quad \text{for } E_n \neq E_m. \quad (6.11)$$

This is called the *orthogonality condition for the eigenfunctions*. Here it should be mentioned that for the same energy eigenvalue E_n , many linearly independent eigenfunctions can exist. The most general solution corresponding to this common value of the energy has, then, the form

$$u_n = c_1 u_{n,1} + c_2 u_{n,2} + \cdots + c_g u_{n,g},$$

where $c_1 \dots c_g$ are arbitrary constants. Their number g , giving the *degree of degeneracy* of the state, is equal to the maximum number of linearly independent solutions that exist for this state. (In the above-mentioned example of the hydrogen

⁶ See A. Sommerfeld, "Wave Mechanics" Dutton, New York, (1929); H.B.G. Casimir, Rotation of a rigid body in quantum mechanics, Thesis (1931); F. Hund, Z. Physik, 51, 11 (1938).

⁷ The more general case of the symmetric top is discussed, e.g., in A.R. Edmonds, Angular Momentum in Quantum Mechanics, Princeton University Press, (1957), pp. 65-67.

It is worth drawing attention in this connection to the work of Flügge and Weiguny: S. Flügge and A. Weiguny, Z. Physik, 171, 171 (1963) where the quantum mechanical rigid body problem is handled carefully.

* Here it is assumed that E is real. This follows, however, from eq. (6.4) with $v = u = u_n$ since this relation yields

$$E_n \int u_n^* u_n dV = E_n^* \int u_n u_n^* dV; \text{ i.e., } E_n^* = E_n.$$

atom* the state n has the degeneracy⁸ n^2 and for a planar oscillator** the degeneracy is $n + 1$; the states of the linear harmonic oscillator are non-degenerate. We can always orthogonalise the basis $u_{n,1}, \dots, u_{n,g}$ of solutions of eq. (6.3) for a given E_n , by taking suitable linear combinations, for which the condition (6.11) is fulfilled for all the different pairs u_n, u_m ; we shall assume this to be the case in what follows. Since each u_n is determined only up to a constant factor, we can assume them to be normalised according to

$$\int u_n^* u_n dq = 1. \quad (6.11')$$

Thus a complex factor of modulus 1 remains undetermined in u_n .

For any arbitrary function f , for which $\int |f|^2 dq$ exists, we have the possibility of a series expansion:

$$f \sim a_1 u_1 + \dots + a_n u_n + \dots, \quad (6.12)$$

where

$$a_n = \int f u_n^* dq. \quad (6.13)$$

The relation (6.13) follows from the orthonormality relations (6.11), (6.11'). The sign \sim denotes that the series (6.12) is not, in general, convergent in the usual sense, but it is convergent only in the mean. That is,

$$\lim_{N \rightarrow \infty} \int \left| f - \sum_{k=1}^N a_k u_k \right|^2 dq = 0. \quad (6.12')$$

This is particularly to be noted if f possesses any square-integrable singularities. The integral in the relation (6.12') can be expanded as

$$\begin{aligned} \int |f|^2 dq &= \sum_{k=1}^N a_k^* \int f u_k^* dq - \sum_{k=1}^N a_k \int f^* u_k dq \\ &\quad + \sum_{i,k=1}^N a_k a_i^* \int u_k u_i^* dq \end{aligned}$$

or taking account of (6.11), (6.11') and (6.13), we have

$$\begin{aligned} \int |f|^2 dq &= 2 \sum_{k=1}^N a_k^* a_k + \sum_{k=1}^N a_k a_k^* \\ &= \int |f|^2 dq - \sum_{k=1}^N a_k^* a_k. \end{aligned}$$

Consequently (6.12') is equivalent to

$$\int |f|^2 dq = \sum_{k=1}^{\infty} a_k^* a_k, \quad (6.14)$$

whereby the convergence of the series occurring in (6.14) is also ensured. This relation is called the *completeness relation*, since this is a criterion that makes sure that no function u_n is missed and no new linearly independent function can be added to f .

*Cf. L.I. Schiff, Quantum Mechanics, Third edition, McGraw-Hill Book Co., (1968), p. 94.

⁸For the additional two-fold degeneracy due to electron spin, see Sec. 13.

**Cf. S. Flügge, loc. cit., problem 42, p. 107.

Since the relation between the function f and the coefficients a_n is linear, it follows, for two arbitrary functions

$$\begin{aligned} f &\sim a_1 u_1 + \cdots + a_n u_n + \cdots; & a_n = \int f u_n^* dq, \\ g &\sim b_1 u_1 + \cdots + b_n u_n + \cdots; & b_n = \int g u_n^* dq, \end{aligned}$$

that

$$\int f^* g dq = \sum_{k=1}^{\infty} a_k^* b_k; \quad \int f g^* dq = \sum_{k=1}^{\infty} a_k b_k^*. \quad (6.14')$$

The last line of (6.14') can be derived if we remember that (6.13) must hold when we substitute for f , $\lambda f + \mu g$ and for a_k , $\lambda a_k + \mu b_k$ where λ and μ are arbitrary numbers.

A generalisation of the orthogonality relations* (6.11) is obtained by introducing the elementary wave-packets $\bar{u}_{\lambda_1 \lambda_2} = \bar{u}_{\lambda_1} - \bar{u}_{\lambda_2}$, defined by eq. (6.6). The application of relation (6.4) to these functions then gives

$$\int \bar{u}_{\lambda_1' \lambda_1''}^* \bar{u}_{\lambda_2' \lambda_2''} dq = 0, \quad (6.11'')$$

if the intervals $(\lambda_1' \lambda_1'')$ and $(\lambda_2' \lambda_2'')$ do not overlap. If further the spectrum is purely continuous, then the limit

$$\lim_{\Delta \lambda \rightarrow 0} \frac{1}{4\Delta \lambda} \int \bar{u}_{\lambda, \lambda + \Delta \lambda}^* \bar{u}_{\lambda, \lambda + \Delta \lambda} dq \rightarrow G(\lambda)$$

exists. Hence

$$\int \bar{u}_{\lambda, \lambda + \Delta \lambda}^* \bar{u}_{\lambda, \lambda + \Delta \lambda} dq = \int_{\lambda}^{\lambda + \Delta \lambda} G(\lambda) d\lambda. \quad (6.15)$$

We can combine (6.11'') and (6.15) into the equation

$$\int \bar{u}_{\lambda_1' \lambda_1''}^* \bar{u}_{\lambda_2' \lambda_2''} dq = \int_{\lambda'}^{\lambda''} G(\lambda) d\lambda, \quad (6.16)$$

where (λ', λ'') represents the sub-interval common to the intervals $(\lambda_1', \lambda_1'')$ and $(\lambda_2', \lambda_2'')$. The functions \bar{u} and u are considered to be *normalised* with respect to the continuous parameter λ , if the function $G(\lambda)$ in (6.16) is, in particular, equal to 1. Thus we have

$$\int \bar{u}_{\lambda_1' \lambda_1''}^* u_{\lambda_2' \lambda_2''} dq = (\lambda'' - \lambda'). \quad (6.17)$$

If we introduce, instead of λ , a function, $\mu = f(\lambda)$ as a new parameter, then for the functions normalised with respect to μ , we have

$$\bar{u}_{\mu, \mu''} = \int_{\lambda'}^{\lambda''} \sqrt{\frac{d\mu}{d\lambda}} d\bar{u}_{\lambda, \lambda}, \quad (6.17')$$

or roughly speaking, the eigenfunctions u_{λ} are to be multiplied by $\sqrt{d\mu/d\lambda}$. In place of the series expansion (6.12) we then have the integral

$$f \sim \int a_{\lambda} u_{\lambda} d\lambda = \int a_{\lambda} d\bar{u}_{\lambda}, \quad (6.18)$$

*For this portion, reference may also be made to D.I. Blokhintsev, Quantum Mechanics, (translated from Russian, fourth edition by J.B. Sykes and M.J. Kearsley), Reidel Pub. Co. Dordrecht, Holland, (1964).

where

$$a_\lambda G(\lambda) = \int f u_\lambda^* dq \quad (6.19)$$

and

$$a_\lambda = \int f u_\lambda^* dq, \quad (6.19')$$

in case u_λ is normalised according to (6.17). The completeness relation now reads as

$$\lim_{\substack{\lambda_1 \rightarrow -\infty \\ \lambda_2 \rightarrow +\infty}} \int \left| f - \int_{\lambda_1}^{\lambda_2} a_\lambda u_\lambda d\lambda \right|^2 dq = 0$$

or

$$\lim_{\substack{\lambda_1 \rightarrow -\infty \\ \lambda_2 \rightarrow +\infty}} \int \left| f - \int_{\lambda_1}^{\lambda_2} a_\lambda d\bar{u}_\lambda \right|^2 dq = 0,$$

which, from the orthogonality relations, is equivalent to

$$\int |f|^2 dq = \int |a_\lambda|^2 d\lambda. \quad (6.20)$$

If the function $G(\lambda)$ is not normalised, we would have

$$\int |f|^2 dq = \int |a_\lambda|^2 G(\lambda) d\lambda. \quad (6.20')$$

Between two functions f and g and their expansion co-efficients a_λ and b_λ , respectively there exists the relation

$$\int f g^* dq = \int a_\lambda b_\lambda^* G(\lambda) d\lambda. \quad (6.20'')$$

We can also formulate these results such that the discrete spectrum (also sometimes called point spectrum) and the continuous spectrum can be given a uniform treatment.⁹ Let \bar{u}_λ again be the functions defined in (6.6') which change with λ discontinuously; further let us again set, for brevity,

$$\bar{u}_{\lambda' \lambda''} = \bar{u}_{\lambda''} - \bar{u}_{\lambda'}.$$

Further (λ', λ'') is the interval common to the intervals $(\lambda'_1, \lambda''_1)$ and $(\lambda'_2, \lambda''_2)$ in case both these intervals are relevant to the problem. Sometimes such a common interval may not exist at all. Without normalisation, we write the orthogonality relation

$$\int \bar{u}_{\lambda'_1 \lambda''_1}^* \bar{u}_{\lambda'_1 \lambda''_1} dq = \bar{G}(\lambda'') - \bar{G}(\lambda'), \quad (6.16')$$

where $\bar{G}(\lambda)$ is a non-decreasing, positive definite function of λ which may have jump discontinuities. Waiving the normalisation condition has the advantage that we can do the calculation better for the degenerate case, where we have to consider a linear combination of eigenfunctions belonging to a particular value λ instead of a single eigenfunction \bar{u}_λ . If we write instead of (6.19) the relation

$$\int_{\lambda'}^{\lambda''} a_\lambda d\bar{G}(\lambda) = \int f \bar{u}_{\lambda' \lambda''}^* dq, \quad (6.19'')$$

then this holds for discrete as well as for continuous spectra. The completeness relation can then be written as

$$\int |f|^2 dq = \int_{-\infty}^{+\infty} |a_\lambda|^2 d\bar{G}_\lambda \quad (6.21)$$

and

$$\int f g^* dq = \int_{-\infty}^{+\infty} a_\lambda b_\lambda^* d\bar{G}_\lambda. \quad (6.22)$$

Now we consider the operator P_λ , which associates the arbitrary function f with the portion of the corresponding integral in (6.18) cut off at λ :

$$P_\lambda f = \int_{-\infty}^{\lambda} a_\lambda d\bar{u}_\lambda. \quad (6.22')$$

⁹ Cf. for this J.v. Neumann, Göttinger Nachr. (1927), 1.

This operator is called a *projection operator*, since it projects the manifold of a_λ onto a sub-manifold – the one for which a_λ vanishes outside the interval $(-\infty, \lambda)$. Obviously the projection operator P has the property

$$P^2 = P, \quad (6.23)$$

and conversely we shall call every operator P with this property, a projection operator. In our case, the operator

$$P_{\lambda''\lambda'} = P_{\lambda''} - P_{\lambda'}$$

is certainly a projection operator for $\lambda'' > \lambda'$ because

$$(P_{\lambda''\lambda'})^2 = (P_{\lambda''} - P_{\lambda'})^2 = P_{\lambda''\lambda'} \quad \text{for all } \lambda'' > \lambda'. \quad (I)^*$$

Further

$$P(-\infty) = 0; \quad P(+\infty) = 1. \quad (II)$$

For $\lambda' > \lambda$ we have $P_{\lambda'} \rightarrow P_{\lambda''}$ when $\lambda' \rightarrow \lambda^+$. Now we seek the relationship between the operator H and the operator P_λ . According to (6.3')

$$H\bar{u}_{\lambda''\lambda'} = \int_{\lambda'}^{\lambda''} E(\lambda) d\bar{u}_\lambda$$

Hence

$$P_{\lambda''\lambda'} f = \int_{\lambda'}^{\lambda''} a_\lambda d\bar{u}_\lambda; \quad H P_{\lambda''\lambda'} f = \int_{\lambda'}^{\lambda''} a_\lambda E(\lambda) d\bar{u}_\lambda = \int_{\lambda'}^{\lambda''} E(\lambda) d(P_{\lambda''\lambda'} f). \quad (6.24)$$

If $\lambda' = -\infty, \lambda'' = +\infty$, this last equation transforms into

$$Hf = \int_{-\infty}^{+\infty} E(\lambda) d(P_\lambda f), \quad (III)$$

which is equivalent to the statement that for all g we have

$$\int (g^* H f) dq = \int_{-\infty}^{+\infty} E(\lambda) d \left(\int g^* P_\lambda / dq \right). \quad (III')$$

Instead of the arbitrary parameter λ we can introduce the energy itself. The requirements (I), (II), and (III) define the eigenvalue problem in a very general manner. We shall deal with the solution of this problem in a later section.

For future calculations it is convenient to introduce the integral

$$\int u_\lambda^* u_\lambda / dq$$

as an improper integral: Let $\delta(\lambda)$ be an improper function** with the property that for all continuous functions $f(\lambda)$ we have

$$\int_{\lambda_1}^{\lambda_2} f(\lambda) \delta(\lambda) d\lambda = \begin{cases} f(0) & \text{if } 0 \text{ is inside } (\lambda_1 \lambda_2), \\ 0 & \text{if } 0 \text{ is outside } (\lambda_1 \lambda_2). \end{cases} \quad (6.25)$$

Then we have

$$\int u_\lambda^* u_\lambda / dq = G(\lambda) \delta(\lambda' - \lambda) \quad \text{or} \quad = \delta(\lambda' - \lambda). \quad (6.26)$$

Later we shall also need the derivative δ' , which is defined by

$$\int f(\lambda) \delta'(\lambda) d\lambda = - \int f'(\lambda) \delta(\lambda) d\lambda = \begin{cases} -f'(0) & \text{if } 0 \text{ is inside } (\lambda_1 \lambda_2), \\ 0 & \text{if } 0 \text{ is outside } (\lambda_1 \lambda_2). \end{cases} \quad (6.27)$$

These relations are to be considered as convenient abbreviations for (6.16) and (6.17).

*This can be verified by using (6.22') and (6.23).

**This is, of course, the Dirac δ -function around which much mathematical literature has developed over the years. Cf. P.A.M. Dirac, The Principles of Quantum Mechanics, fourth edition, Oxford University Press, 1959, p. 58 and ff. See also M.J. Lighthill, Introduction to Fourier Analysis and Generalised Functions, Cambridge University Press, (1964).

It may be noted that the eigenvalue problem for the general differential equation (6.3) is always equivalent¹⁰ to a variational problem

$$\delta \int u^* (\mathbf{H} u) dq = \delta \int (\mathbf{H} u)^* u dq = 0 \quad (6.28)$$

with the auxiliary condition

$$\int u^* u dq = 1. \quad (6.29)$$

Here u and u^* are to be varied independently of each other. Under certain circumstances, eq. (6.28) can be rearranged using integration by parts. For example, in cartesian co-ordinates and for the Hamiltonian

$$\mathbf{H} = - \sum_k \frac{\hbar^2}{2m_k} \left(\frac{\partial}{\partial q_k} - \frac{i\epsilon_k}{\hbar c} \Phi_k \right)^2 + V(q),$$

we have the variational principle

$$\delta \int \left[\sum_k \frac{\hbar^2}{2m_k} \left(\frac{\partial u^*}{\partial q_k} + \frac{i\epsilon_k}{\hbar c} \Phi_k \right) \left(\frac{\partial u}{\partial q_k} - \frac{i\epsilon_k}{\hbar c} \Phi_k \right) + V u^* u \right] dq = 0 \quad (6.30)$$

with the subsidiary condition (6.29). On account of eq. (6.3), the value of the integral (6.30) for the extremal function is always equal to the energy eigenvalue. This variational problem is often useful for the approximate integration of differential equations* in which one imposes special restrictions on the function u and tries to solve the problem under these additional conditions. The values of the variational integrals are always found to be larger than the true eigenvalues.

In the case of real eigenfunctions, two general laws have yet to be mentioned: (1) The eigenfunction without nodes (null points), if it exists, belongs to the "smallest eigenvalue. (2) In the one-variable case the arrangement of the eigenfunctions according to increasing nodal number, always corresponds to increasing eigenvalues. The number of nodes present turns out to be a "quantum number" of the system.

¹⁰ E. Schrödinger, Ann. d. Phys. 79, 361 (1926).

* Cf. Courant and Hilbert, loc. cit., Ch. 4. The application of variational methods to quantum mechanical problems is not discussed further in this book. A useful reference in this regard is "The Variational Method in Quantum Theory" by B.L. Moiseiwitsch, edited by D.R. Bates, Academic Press, New York, (1961), vol. I, pp. 211-228.

CHAPTER IV

Matrix Mechanics

7. General Transformations of Operators and Matrices

We can establish an important connection between the operators acting on the functions u_n and the matrices associated with the operators, using the completeness relation. If F is a linear operator, then corresponding to every eigenfunction u_n there is an expansion

$$(Fu_n) \sim \sum_k u_k F_{kn} \quad \text{with} \quad F_{kn} = \int u_k^* (Fu_n) dq. \quad (7.1)$$

When F is Hermitian,

$$\int u_k^* (Fu_n) dq = \int (Fu_k)^* u_n dq,$$

and hence

$$F_{kn} = (F_{nk})^*, \quad (7.2)$$

implying that the matrix is Hermitian. Next let us consider two Hermitian operators F and G :

$$Fu_n \sim \sum_k u_k F_{kn}; \quad F_{kn} = \int u_k^* (Fu_n) dq,$$

$$Gu_m \sim \sum_k u_k G_{km}; \quad G_{km} = \int u_k^* (Gu_m) dq.$$

Applying the completeness relation (6.14') (valid for all n, m), which states that

$$\int (Fu_n)^* (Gu_m) dq = \sum_{k=1}^{\infty} F_{kn}^* G_{km}$$

and using the Hermitian property of F , we have

$$\int u_n^* (FGu_m) dq = \sum_k F_{nk} G_{km}. \quad (7.3)$$

On the basis of eq. (7.1), a matrix can be associated with every operator. In particular to every Hermitian operator there is associated a Hermitian matrix. From eq. (7.3) then follows the law: *With the product $F \cdot G$ of two operators is associated the product $(F) \cdot (G)$ of the corresponding matrices, if the latter product is formed according to the usual rule for matrix multiplication.* As is well known, the rule for multiplication of matrices reads,

$$(FG)_{nm} = \sum_k F_{nk} G_{km}. \quad (7.3')$$

Notice that here we have only assumed the Hermiticity of F and G and not that of

$(F \cdot G)$. No assumption about the commutability of F and G has also been made.

If now any arbitrary function f is expanded as $\sum a_k u_k$, then to Ff corresponds the expansion $\sum_k (Fu_k) a_k = \sum_{k,l} u_l F_{lk} a_k$.¹ Hence the operator F connects the coefficients $a_1, a_2, \dots, a_m, \dots$ and $b_1, b_2, \dots, b_n, \dots$ through the relation

$$b_n = \sum_m F_{nm} a_m. \quad (7.4)$$

Thus the matrix, equivalent to the operator F acts as a *linear transformation* in the infinite dimensional vector-space of the expansion coefficients (a_m) of the system of functions u_n . The Hermiticity of F is expressed by the relation

$$\sum_n a_n^* b_n = \sum_n a_n b_n^*. \quad (7.5)$$

This "scalar product" is, therefore, real.

From the general expression (7.1) the diagonal elements of the matrix can be written as

$$F_{nn} = \int u_n^* (Fu_n) dq. \quad (7.6)$$

We have seen in the previous chapter that in simple cases this expression has the significance of a mean or expectation value. To obtain this result, we started from the expression $|\phi(p)|^2 dp$ which gives the probability that the momentum of the state of the system considered lies between p and $p + dp$. Then the mean value of an arbitrary rational integral function, F of p , is given by

$$\int F(p) |\phi(p)|^2 dp = \int \psi^*(q) \left[F\left(\frac{\hbar}{i} \frac{\partial}{\partial q}\right) \psi(q) \right] dq.$$

Similarly for an arbitrary function of q we have

$$\bar{F}(q) = \int \psi^* F(q) \psi dq,$$

and this agrees with (7.6), if we interpret the action of the operator F in (7.6) on the wave function ψ as mere multiplication by the function $F(q)$. If the function F depends linearly on p and arbitrarily on q , i.e.

$$F = \frac{1}{i} \sum_k [A_k(q) p_k + p_k A_k(q)]$$

(note that symmetrisation of the expression on the right-hand side is required for F to be Hermitian), we can introduce the current density i_k corresponding to the k^{th} degree of freedom in Cartesian co-ordinates and thus obtain

$$\bar{F} = \sum_k \int A_k(q) \left[m i_k + \frac{e}{c} \Phi_k \right] dV = \sum_k \int \psi^* \frac{1}{2} \left[A_k \frac{\hbar}{i} \frac{\partial \psi}{\partial q_k} + \frac{\hbar}{i} \frac{\partial}{\partial q_k} (A_k \psi) \right] dq.$$

¹ Strictly speaking, this needs further justification, since the expansion $\sum a_k u_k$ need not converge, but need only agree with f in the sense of convergence in the mean. If we denote the N^{th} partial sum $\sum a_k u_k$ by f_N , this means that

$$\lim_{N \rightarrow \infty} \int |f - f_N|^2 dq = 0$$

We now demand that the operator F has the following property: If there is a function g corresponding to a sequence g_N defined through $Fg_N = g_N$ such that $\lim_{N \rightarrow \infty} \int |g - g_N|^2 dq = 0$, then $Ff = g$. Two functions g and g' , for which $\int |g - g'|^2 dq = 0$, are not considered to be basically different. In our case the existence of such a function is assured, if the sum $\sum_k |b_k|^2$ converges ($b_k = \sum_i F_{ki} a_i$) and this is again true if the sums

$$\sum_i |F_{ki}|^2 = \sum_i |F_{ik}|^2$$

converge for every k .

Since the expectation value of the sum of two quantities is equal to the sum of their expectation values and (7.6) is linear in the operator F , (7.6) can be considered to be valid even for sums of operators of the type considered above. In particular this is true if F represents the Hamiltonian H .

Thus, we can state that for any operator F , which can be associated in the above manner with physical quantities, (i.e. with properties of the system which can be described by numerical data that can be uniquely determined ("measured"), in principle, by suitable experimental arrangements) *the diagonal element F_{nn} is equal to the expectation value of the associated quantity F in the state of the system characterised by the eigenfunction u_n* . What the physical quantities of a system are and how they are to be measured, can be decided only by a further development of the theory based on experience.² We further note that even for such a special operator F , *only the diagonal elements of the associated matrix F have a direct physical meaning*. The non-diagonal elements are connected only indirectly with the data obtainable from measurement and that too in the following way: The mean value of an arbitrary, rational integral function $f(F)$ of the operator F , and in particular, of all its powers, are given by the diagonal element $[f(F)]_{nn}$. If, on the other hand, we assume for the formal construction of $[f(F)]_{nn}$ the multiplication law for matrices, we can show that given the diagonal elements, the non-diagonal elements, F_{nm} can be, in general, uniquely determined.

The connection between matrices and operators expressed through the relation (7.1) presupposes only that the functions $u_1, u_2, \dots, u_n, \dots$ form a complete orthogonal set. It is worthwhile to investigate as to how the matrices themselves change, if one goes from this set to a new set $v_1, v_2, \dots, v_n, \dots$, which also possesses the properties of orthogonality and completeness. Each of the functions v_m has the expansion

$$v_m \sim \sum_n u_n S_{nm} \text{ with } S_{nm} = \int u_n^* [S u_m] dq = \int u_n^* v_m dq. \quad (7.7)$$

The matrix S_{nm} defines a linear operator S , which associates with every function $f = a_1 u_1 + a_2 u_2 + \dots$ the function $g = a_1 v_1 + a_2 v_2 + \dots$, the expansion coefficients being the same for the two sets. The operator S has, therefore, the property

$$\int f u_n^* dq = \int [S f] v_n^* dq$$

for all f . The operator S , however, is basically different from the Hermitian operators F , considered earlier, which may correspond to physical quantities. We call S a *transformation operator*. Since according to the completeness relation, we have

$$\int v_n^* v_m dq = \sum_k S_{kn}^* S_{km},$$

the orthogonality and normalisation of the v_n are expressed by

$$\sum_k S_{kn}^* S_{km} = \begin{cases} 0 & \text{for } n \neq m, \\ 1 & \text{for } n = m. \end{cases} \quad (7.8)$$

² This standpoint is opposed to the point of view that to every Hermitian operator corresponds a quantity or "observable" of the system and that it is always meaningful to speak of the probability that this quantity F has a determinate value in the state concerned (cf. Sec. 9).*

* This point is further discussed, for example, by Jauch who gives a simple counter-example: J.M. Jauch in "Aspects of Quantum Mechanics" edited by Abdus Salam and E.P. Wigner, Cambridge University Press (1972), pp. 151, 152.

Since from $f \sim \sum a_k u_k$, $g \sim \sum a_k v_k$, we have $\sum_k |a_k|^2 = \int |f|^2 dq = \int |g|^2 dq$; this is also equivalent to the requirement that for all f we should have

$$\int |Sf|^2 dq = \int |f|^2 dq. \quad (7.9)$$

Hence for two arbitrary functions f and g , we get

$$\int (Sf)^* (Sg) dq = \int f^* g dq. \quad (7.9')$$

The operator S is, therefore, not Hermitian, but has according to (7.9), the property of "preserving length", if we take $\int |f|^2 dq$ as the measure of the "length" of a function $\int |f - g|^2 dq$ as the measure of the "distance" between two functions, f and g . If further $v_n = (Su_n)$ are to form a complete set of functions, then the manifold of Sf must coincide with that of f . That is, *the operator S^{-1} which is the inverse of S must exist for all f* . The completeness of the set of functions v implies that we can expand u_n also in terms of v_n according to

$$u_m \sim \sum_n v_n S_{nm}^{-1}; \quad S_{nm}^{-1} = \int v_n^* [S^{-1} v_m] dq = \int v_n^* u_m dq. \quad (7.7')$$

By comparison with (7.7) it follows that

$$S_{nm}^{-1} = S_{mn}^*. \quad (7.10)$$

Thus S is not a Hermitian matrix. The Hermitian conjugate, \tilde{S} of S obtained from the matrix S by taking the complex conjugate and transpose is identical with the matrix inverse to S , i.e.

$$S^{-1} = \tilde{S}. \quad (7.11)$$

In fact, on account of $S_{kn}^* = \tilde{S}_{nk}$, eq. (7.8) is the same as

$$\tilde{S} S = I, \quad (7.8')$$

where I denotes the unit matrix. The completeness of the set of functions v_n now implies that

$$\sum_k (S_{kn}^{-1})^* S_{km}^{-1} = \begin{cases} 0 & \text{for } n \neq m, \\ 1 & \text{for } n = m, \end{cases} \quad (7.12)$$

as can be seen by applying the completeness relation to (7.7'). Using eq. (7.10), we can write eq. (7.12) also as

$$\sum_k S_{nk} S_{mk}^* = \begin{cases} 0 & \text{for } n \neq m, \\ 1 & \text{for } n = m, \end{cases} \quad (7.13)$$

or in matrix notation,

$$S \tilde{S} = I. \quad (7.13')$$

It is important to stress that (7.13') does not follow immediately from (7.8'), but does so if and only if the inverse operator S^{-1} exists for all f . An operator S is called *unitary* if it possesses the inverse operator, besides having the property of preserving length [eq. (7.9)]; we call the corresponding matrix S *unitary*, if it satisfies both the relations (7.8') and (7.13'). The transformation operators are unitary operators of this type. The composition (multiplication) of two unitary operators (matrices) obviously generates again a unitary operator (a unitary matrix).

It is perhaps useful to stress the similarity of the above relations to similar ones for matrices of finite order. Here the complete set of functions consists of N orthogonal (complex) vectors $u_k(1), u_k(2), \dots, u_k(N)$, in which the arguments are the values q_1, \dots, q_N of the concerned quantities, instead of q .

An orthogonal system of vectors in a finite-dimensional space is complete if and only if the number p of the vectors of the system coincides with the dimension N of the system. The case, where a matrix S satisfies the relation

$$\tilde{S}S = I$$

but $S\tilde{S} \neq I$ arises if S is not a square but a rectangular matrix, S_{nm} ($n = 1, 2, \dots, N$; $m = 1, 2, \dots, p$), the column number p of which is smaller than the row number N . In this case it is clear that $S\tilde{S}$ is not a unit matrix. So the inverse mapping S^{-1} which takes the vector $f(r) = \sum_{k=1}^p a_k u_k(r); r = 1, \dots, N$ into $(Sf(r)) = \sum_{k=1}^p a_k v_k(r)$, exists for all $f(r)$, only if $p = N$. The requirement that we should also have $S\tilde{S} = I$ besides $\tilde{S}S = I$ is the same as saying that S should be a square matrix. For infinite dimensional matrices, i.e. matrices with an infinite number of rows and columns, however, the completeness of the set cannot be determined by a simple counting, and hence, besides the relation $S\tilde{S} = I$, we must also require the unrestricted existence of the reciprocal.

We can now answer the question as to how the matrix associated with the operator F according to (7.1) changes when we go over from the set of functions u_k to the set v_k . We have, according to (7.7),

$$\left. \begin{aligned} F'_{nm} &= \int v_n^* (F v_m) dq, \\ F'_{nm} &= \sum_k \sum_l \int S_{kn}^* u_k^* S_{lm} F(u_l) dq. \end{aligned} \right\} \quad (7.14)$$

Therefore,³

$$\left. \begin{aligned} F_{kl} &= \int u_k^* F(u_l) dq, \\ F'_{nm} &= \sum_{k,l} S_{kn}^* F_{kl} S_{lm} \end{aligned} \right\} \quad (7.15)$$

or in matrix form

$$(F') = \tilde{S} F S = S^{-1} F S. \quad (7.15a)$$

On inverting, we have

$$(F) = S F' S^{-1} = S F' \tilde{S} \quad (7.15b)$$

We easily verify that F' is a Hermitian matrix if F is a Hermitian matrix. In operator language, eq. (7.15) evidently means: If F transforms the function f into the function g , then F' transforms the function Sf into the function Sg . In this interpretation, however, the "co-ordinate system" of functions u_1, \dots, u_n is considered to be fixed and the operator changes, whereas we started with the fixed operator F and allowed the co-ordinate system to change.

The fundamental C.R. (5.21) for the operators p_k, q_l obviously transform into the corresponding matrix equations, if we replace the function f on which these operators act by the sequence $a_1, a_2, \dots, a_n, \dots$ of the coefficients of expansion of f in terms of an arbitrary, complete, orthogonal set $v_1, v_2, \dots, v_n, \dots$. In fact, these C.R. remain invariant under arbitrary unitary transformations of the form (7.15).

Now we can easily describe the special eigenfunctions u_1, u_2, \dots , which satisfy the equation

$$H u = E u$$

and which correspond to the stationary states of our system with the Hamiltonian operator H . The functions u are orthonormal. For these functions the matrix corresponding to the Hamiltonian (energy), defined by (7.1), viz.,

$$(H_{nm}) = \int u_n^* (H u_m) dq$$

³ In this derivation we have refrained from dealing with questions of convergence. The problem of convergence of the sums appearing in (7.15) is, in general, complicated.

is a *diagonal matrix*:

$$H_{nm} = \begin{cases} 0 & \text{for } n \neq m, \\ E_n & \text{for } n = m. \end{cases} \quad (7.16)$$

Using this result we can also formulate the problem of the determination of the energy eigenvalues E_n and the matrices $p_{n,m}^{(k)}$, $q_{n,m}^{(k)}$ without any knowledge of the wave equation. We first write down the C.R. using the rule for multiplying matrices and then impose the condition that the energy matrix, when expressed in terms of the matrix elements of $p^{(k)}$ and $q^{(k)}$ with the help of this rule, should be diagonal. We obtain, in this way, an infinite set of equations to determine the matrix elements of p and q as well as the energy eigenvalues. This was the basis of the "Matrix mechanics" of Heisenberg⁴ which came, historically speaking, before Wave Mechanics. These equations can in practice be solved only in a few cases, e.g., for a harmonic oscillator. In the case of the hydrogen atom we can obtain the energy values (by this method) but the evaluation of the matrix elements of p and q is no longer possible.* This is connected with the fact that a continuous spectrum is also present, in which case the matrix calculation cannot be done. We shall come back to this point presently. On the other hand, the matrix calculus is convenient if we are dealing with a sub-space of finite dimensions. Such sub-spaces occur, for example, in degenerate systems, i.e. if there are many states (say, g) corresponding to a particular energy value. In this case we can make, within this g -dimensional sub-space, an arbitrary unitary transformation of the matrices according to (7.15), without disturbing the equations which are to be solved. This is because in this sub-space the energy matrix will be equal to the unit matrix multiplied by a fixed number E . Then this matrix remains invariant under the transformation considered. In Wave Mechanics this corresponds to the possibility of transforming orthonormal eigenfunctions u_1, \dots, u_g belonging to E according to

$$u'_m = \sum_{n=1}^g u_n S_{nm} \quad (m = 1, 2, \dots, g).$$

If, in particular, S is a *unitary* (square) matrix of g -rows, then the orthogonality of the set of functions remains unaltered under this transformation. In matrix language, the eigenvalue problem can also be posed in a slightly different way. We substitute for the p_k and q_k certain matrices which satisfy the C.R. (wave-mechanically, they can correspond to any orthogonal set of functions). On using the multiplication rule for matrices, we obtain for the energy, a certain Hermitian matrix H_{mn} . The task now is to bring this matrix to diagonal form by unitary transformations, i.e. to solve the infinite number of linear equations

$$S^{-1} H S = E$$

or

$$HS = SE; \quad \sum_m H_{nm} S_m = S_n E \quad (7.17)$$

⁴ W. Heisenberg, Z. Physik 33, 879 (1925).

*This calculation was done by Pauli himself: W. Pauli, Jr. Z. Phys., 36, 336 (1926).

For other applications of matrix mechanics, including perturbation calculations in the Heisenberg picture, see Max Born and P. Jordan, Elementare Quantenmechanik, Springer Verlag, (1930); also H.S. Green, Matrix Mechanics, P. Noordhoff Ltd., Netherlands, (1965).

for every possible energy matrix E . The condition that S is unitary imposes on the coefficients S_n the normalisation

$$\sum_n |S_n|^2 = 1. \quad (7.17')$$

For every possible value E_p of E we thus obtain a system of coefficients $S_{n,p}$ and it is easy to see that this will be unitary on account of eq. (7.8), in case H is Hermitian. Though this system of equations cannot in general be solved, we shall see that they can be used in perturbation theory.⁵

We shall now discuss the generalisation of these results to continuous spectra, disregarding the question of convergence. If the eigenfunctions u are normalised with reference to the parameter λ or the parameters λ_1, \dots , then we can form, analogous to (7.1), the expressions

$$F u_\lambda \sim \int u_\lambda F_{\lambda' \lambda} d\lambda'; \quad F_{\lambda' \lambda} = \int u_{\lambda'}^* (F u_\lambda) dq. \quad (7.18)$$

The $F_{\lambda' \lambda}$ and in fact u itself may contain the δ -function defined by (6.25) and (6.27). For example, we can choose for λ , the numerical value of q and for u the δ -function, i.e. $u_q(q) = \delta(q - q')$, where $\delta(q - q')$ is an abbreviation for $\delta(q_1 - q'_1)\delta(q_2 - q'_2) \dots \delta(q_f - q'_f)$. For, we then have

$$\int u_q^*(q) u_{q'}(q) dq = \int \delta(q - q') \delta(q - q'') dq = \delta(q' - q'').$$

The matrix of q will be

$$q_k(q', q'') = \int \delta(q - q') q_k \delta(q - q'') dq$$

or

$$q_k(q', q'') = q'_k \delta(q' - q''). \quad (7.19a)$$

Similarly

$$p_k(q', q'') = \int \delta(q - q') \frac{\hbar}{i} \frac{\partial}{\partial q_k} \delta(q - q'') dq$$

or

$$p_k(q', q'') = \frac{\hbar}{i} \delta'_k(q' - q''), \quad (7.19b)$$

if the latter expression is written as an abbreviation for

$$\delta(q'_1 - q''_1) \dots \delta(q'_{k-1} - q''_{k-1}) \delta'(q'_k - q''_k) \delta(q'_{k+1} - q''_{k+1}) \dots \delta(q'_f - q''_f).$$

We can verify the formal validity of

$$\int [p_k(q', q'') q_k(q''', q'') - q_k(q', q'') p_k(q''', q'')] dq''' = \frac{\hbar}{i} \delta(q' - q'').$$

Actually, these symbols have a meaning only if they are considered to be multiplied by arbitrary functions, f^* and g , which are functions of q' and q'' respectively and integrated over them, e.g.

$$\int_{q'_1}^{q''_1} \int_{q'_2}^{q''_2} f^*(q') q_k(q', q'') g(q'') dq' dq'' = \int_{q'_1}^{q''_1} f^*(q) q_k g(q) dq.$$

Here (q_1, q_2) represents the sub-interval common to the intervals (q'_1, q'_2) and (q''_1, q''_2) .

Analogous to (7.7) we have the transformation of a complete set u_λ to another set v_μ , according to

$$v(\mu; q) = \int u(\lambda; q) S(\lambda, \mu) d\lambda$$

with

$$S(\lambda, \mu) = \int u^*(\lambda; q) v(\mu; q) dq. \quad (7.20)$$

⁵ Cf. M. Born, P. Jordan and W. Heisenberg, Z. Physik, 35, 557 (1926).

Here $S(\lambda, \mu)$ is a unitary matrix which satisfies relations similar to (7.8) and (7.13)

$$\int S^*(\lambda, \mu) S(\lambda, \mu') d\lambda = \delta(\mu - \mu'), \quad (7.21)$$

$$\int S(\lambda, \mu) S^*(\lambda', \mu) d\mu = \delta(\lambda' - \lambda). \quad (7.22)$$

The transformation formula, which is similar to (7.15) reads

$$F(\mu; \mu') = \int S^*(\lambda, \mu) F(\lambda, \lambda') S(\lambda', \mu') d\lambda d\lambda'. \quad (7.23)$$

Of particular interest is the case when one of the indices is discrete and the other continuous. In this case, we have

$$v_n(q) = \int u(\lambda; q) S_n(\lambda) d\lambda; \quad u(\lambda; q) = \sum_n v_n^*(q) S_n^*(\lambda), \quad (7.24)$$

$$S_n(\lambda) = \int u^*(\lambda; q) v_n(q) dq. \quad (7.25)$$

$$\int S_n^*(\lambda) S_m(\lambda) d\lambda = \delta_{n,m}; \quad \sum_n S_n(\lambda) S_n^*(\lambda') = \delta(\lambda - \lambda'), \quad (7.26)$$

$$(\delta_{n,m} = 0 \text{ for } n \neq m; \quad \delta_{n,m} = 1 \text{ for } n = m)$$

$$F_{n,m} = \int S_n^*(\lambda) F(\lambda, \lambda') S_m(\lambda') d\lambda d\lambda'; \quad F(\lambda, \lambda') = \sum_{n,m} S_n(\lambda) F_{n,m} S_m^*(\lambda'). \quad (7.27)$$

These formulae will become particularly simple, if we substitute for $u(\lambda, q)$ the set $u_q(q) = \delta(q - q')$. Then we have, evidently

$$S_n(q') = \int \delta(q - q') v_n(q) dq = v_n(q'), \quad (7.25')$$

$$\int v_n^*(q) v_m(q) dq = \delta_{n,m}; \quad \sum_n v_n(q) v_n^*(q') = \delta(q - q'), \quad (7.26')$$

$$F_{n,m} = \int v_n^*(q) F(q, q') v_m(q') dq dq'; \quad F(q, q') = \sum_{n,m} v_n(q) F_{n,m} v_m^*(q'). \quad (7.27')$$

We obtain the important result that the eigenfunctions $\bar{v}_n(q)$ themselves appear as special transformation functions for going from the set $\delta(q - q')$ to the set $\bar{v}_n(q)$. If we choose, in particular, for $\bar{v}_n(q)$ the set $u_n(q)$ that converts a given Hamiltonian operator H to diagonal form, $H_{n,m} = E_n \delta_{n,m}$, we can say that according to (7.25'), the functions $u_n(q)$ are the *transformation functions from the operator q to the operator H*. For, q will be brought to diagonal form by the set of δ -functions, $\delta(q - q')$, and H by the set of function $u_n(q)$. As regards the second relation in (7.26'), we note that it is only a symbolic expression for the completeness relation. By integration we arrive at an "effective" equation involving two arbitrary functions f and g :

$$\sum_n \int f^*(q) v_n(q) dq \cdot \int g(q') v_n^*(q') dq' = \int f^* g dq.$$

This is in exact agreement with the completeness relation (6.14'). In particular this equation is correct if f and g are each equal to 1 within some definite regions (which need not be the same) and vanish outside them.

8. The General Form of the Laws of Motion

We have introduced as the basis of matrix mechanics, in addition to the C.R., the prescription that the energy matrix, corresponding to eq. (7.16), should be in the diagonal form:

$$H_{n,m} = E_n \delta_{n,m}. \quad (8.1)$$

Heisenberg has added a further condition concerning the dependence of the matrix elements on time. He stipulated that the time dependence of matrix elements of quantities, which do not explicitly contain the time variable, should be given by

$$F_{n,m}(t) = F_{n,m}(0) e^{\frac{i}{\hbar}(E_n - E_m)t} \quad (8.2)$$

On introducing the unitary diagonal matrix,

$$U_{n,m}(t) = \delta_{n,m} e^{\frac{i}{\hbar}E_n t}, \quad (8.3)$$

we should have

$$\mathbf{F}(t) = \mathbf{U}(t) \mathbf{F}(0) \mathbf{U}^{-1}(t) = \mathbf{U}(t) \mathbf{F}(0) \tilde{\mathbf{U}}(t). \quad (8.4)$$

If the Hamiltonian does not contain time explicitly, eq. (8.4) with $\mathbf{F} = \mathbf{H}$ indicates that eq. (8.1) remains true for all t . The relation (8.2) can also be replaced by the differential equation

$$\frac{\hbar}{i} \dot{F}_{n,m} = F_{n,m} (E_n - E_m) \quad (8.5)$$

or

$$\frac{\hbar}{i} \dot{\mathbf{F}} = \mathbf{H}\mathbf{F} - \mathbf{F}\mathbf{H}, \quad (8.6)$$

if \mathbf{H} is taken to be the diagonal matrix (8.1). On account of (8.1), the form (8.2) again follows from (8.6).¹

The relation (8.2) is nothing but the relation (7.1) for the calculation of the matrix elements with the functions u_n replaced by the functions

$$\psi_n(t) = u_n e^{\frac{i}{\hbar}E_n t} \quad (8.7)$$

These satisfy the wave equation

$$-\frac{\hbar}{i} \dot{\psi}_n = \mathbf{H} \psi_n$$

because then

$$F_{k,n} = \int \psi_k^* (\mathbf{F} \psi_n) dq = e^{\frac{i}{\hbar}(E_k - E_n)t} \int u_k^* (\mathbf{F} u_n) dq = e^{\frac{i}{\hbar}(E_k - E_n)t} F_{k,n}(0).$$

We can now generalise this idea by introducing an arbitrary orthogonal system φ_n as the basis for the matrices, without their having necessarily to possess the special form (8.7); but we impose the essential requirement that all the functions $\varphi_n(t)$ should satisfy the wave equation

$$-\frac{\hbar}{i} \dot{\varphi}_n = \mathbf{H} \varphi_n. \quad (8.8)$$

It immediately follows from the fact that for all solutions of (8.8) the integral $\int \psi^* \psi dq$ is constant in time that, for any solution of the form $c_n \varphi_n + c_m \varphi_m$ with arbitrary coefficients c_n, c_m , we have

$$\frac{d}{dt} \int \varphi_n^* \varphi_m dq = 0 \quad (8.9)$$

i.e. the orthogonality and normalisation of the set of functions remain unaltered in time. Hence from the definition

$$F_{n,m} = \int \varphi_n^* (\mathbf{F} \varphi_m) dq \quad (8.10)$$

we obtain, on differentiating, the following relation which is valid for any Hermitian

¹ In contrast to the older treatments of matrix mechanics it should be stressed that (8.6) does not follow from (8.1), unless (8.2) is assumed to be a special postulate.

operator F not containing time explicitly:

$$\begin{aligned}\frac{\hbar}{i} \dot{F}_{n,m} &= \int [(\mathbf{H} \varphi_n)^* (\mathbf{F} \varphi_m) - \varphi_n^* (\mathbf{F} \mathbf{H} \varphi_m)] dq, \\ &= \int [\varphi_n^* (\mathbf{H} \mathbf{F} \varphi_m) - \varphi_n^* (\mathbf{F} \mathbf{H} \varphi_m)] dq, \\ &= (\mathbf{H} \mathbf{F} - \mathbf{F} \mathbf{H})_{n,m}.\end{aligned}$$

Thus we recover eq. (8.6)

$$\frac{\hbar}{i} \dot{\mathbf{F}} = \mathbf{H} \mathbf{F} - \mathbf{F} \mathbf{H}, \quad (8.6)$$

but now without any special assumption about the matrices. On account of this rule, every set of C.R. (which are mutually compatible and do not contain time explicitly) remains unchanged in time, if it is satisfied for $t = 0$.²

Now we introduce the condition (which was not necessary till now) that H does not contain time explicitly. Then from (8.6), H is independent of time. Therefore, there is a unitary operator S independent of time, which diagonalises H :

$$S^{-1} H S = E.$$

It is the same operator S which transforms φ_n to the special form ψ_n given by (8.7) according to the equation

$$\psi_m = \sum_n \varphi_n S_{n,m}$$

By further transformations we obtain from (8.3) and (8.4) with this operator S

$$\mathbf{F}(t) = \mathbf{U}(t) \mathbf{F}(0) \mathbf{U}^{-1}(t) \quad (8.4)$$

with

$$\mathbf{U} = S e^{\frac{i}{\hbar} E t} S^{-1},$$

where by $e^{iEt/\hbar}$ we mean the diagonal matrix $\delta_{n,m} e^{iE_n t/\hbar}$. Here U is obviously unitary, since it arises from the multiplication of unitary matrices. On account of

$$S E S^{-1} = H$$

we can also write U as

$$U(t) = e^{\frac{i}{\hbar} H t} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{i}{\hbar} H t \right)^n = \lim_{N \rightarrow \infty} \left(1 + \frac{it}{N\hbar} H \right)^N$$

which follows directly from (8.11) without requiring E to be diagonal. We can verify further, that U is unitary if all the powers of H exist simultaneously for all functions f . Then we have

$$U(t_1) U(t_2) = U(t_1 + t_2)$$

and in particular

$$U(t) U(-t) = U(0) = I,$$

$$U^{-1}(t) = U(-t) = e^{-\frac{i}{\hbar} H t}$$

² In the older literature on quantum mechanics, we often find the operator equation

$$H t - t H = -\frac{\hbar}{i} I,$$

which arises from (8.6) formally by substituting t for F . It is generally not possible, however, to construct a Hermitian operator (e.g. as function of p and q) which satisfies this equation. This is so because, from the C.R. written above, it follows that H possesses continuously all eigenvalues from $-\infty$ to $+\infty$ (cf. Dirac, Quantum Mechanics, First edition (1930), 34 and 56) whereas on the other hand, discrete eigenvalues of H can be present. We, therefore, conclude that the introduction of an operator t is basically forbidden and the time t must necessarily be considered as an ordinary number ("c-number")* in Quantum Mechanics (cf. for this E. Schrödinger, Berl. Ber. (1931) p. 238).**

*As opposed to this, operators are usually called "q-numbers".

**See also P. Carruthers and M.M. Nieto, Rev. Mod. Phys., 40, 411 (1968) for further references.

If, therefore $U(-t)$ exists for the same set of functions as $U(t)$, the condition for the existence of the inverse is satisfied. We then show, by successive application of the Hermiticity condition

$$\int g^*(H^* f) dq = \int (H^* g)^* f dq$$

to all powers $(H)^n$ of H , that

$$\int g^*(Uf) dq = \int (U^{-1}g)^* f dq,$$

or, in particular for $g = Uf$,

$$\int (Uf)^* (Uf) dq = \int f^* f dq,$$

i.e. the length of f is preserved by U . The operator U satisfies the differential equation

$$\frac{\hbar}{i} \dot{U} = HU; \quad \text{and} \quad \frac{\hbar}{i} \dot{U}^{-1} = -U^{-1}H, \quad (8.11)$$

from which we can readily verify (8.6), using (8.4). The meaning of Uf is that U applied to the expansion coefficients a_1, a_2, \dots of f in terms of the fixed set $\phi_1(0), \phi_2(0), \dots$ transforms these coefficients into $a_n(t) = \sum_m U_{n,m}(t)a_m(0)$. We can also choose, in particular, the set $\delta(q - q')$ for $\phi_k(0)$. Then U transforms the function $f(0)$ to $f(t)$, so that f satisfies the wave equation. For $t = 0$, it coincides with the arbitrary function f :

$$f(q, t) = U(t)f(q, 0) = \int U(q, q'; t)f(q', 0) dq', \quad (8.12)$$

$$U(q, q'; 0) = \delta(q - q'); \quad \frac{\hbar}{i} \frac{\partial}{\partial t} U(q, q'; t) = HU(q, q'; t). \quad (8.13)$$

The fact that U preserves length follows directly from the continuity equation and is implied by

$$\int U(q, q'; t) U^*(q, q''; t) dq = \delta(q' - q''). \quad (8.14)$$

We must also have

$$U^*(q', q; t) = U(q, q'; -t) = U^{-1}(q, q'; t) \quad (8.15)$$

and in general

$$\int U(q, q'; t_0) U(q', q''; t_1) dq' = U(q, q''; t_1 + t_0). \quad (8.16)$$

We recognise this function U as a generalisation of the function $U(x, t)$ appearing in eq. (3.37).³

$$U(q, q'; t) = U(q - q'; t) = e^{-\frac{iq}{\hbar}} \sqrt{\frac{m}{2\pi\hbar}} \frac{1}{\sqrt{t}} e^{\frac{im}{2\hbar} \frac{(q-q')^2}{t}}$$

for the free particle in one dimension. The U -functions for a free particle in three-dimensions and that for an arbitrary number of particles follow on forming the product according to eq. (3.38). For particles which are not free, U will not depend only on the difference $(q - q')$ of the co-ordinates. Further it can in general be constructed only with the help of the eigensolutions $u_n(q)$. For,

$$U(q, q'; t) = \sum_n u_n^*(q') e^{\frac{i}{\hbar} E_n t} u_n(q), \quad (8.17)$$

which satisfies the condition (8.13) on account of (7.26'). The eigenfunctions u_n are

³ The further generalisation of this relation to a continuous eigenvalue spectrum is straight-forward.

characterised by

$$U u_n = e^{\frac{i}{\hbar} E_n t} u_n. \quad (8.18)$$

The existence of a unitary operator U of the type considered here must necessarily be postulated on physical grounds. For, it merely implies that the wave equation can be solved for an arbitrary initial function $f(q, 0)$ and that at time t any function $f(q, t)$ can be reached. This statement also indicates the way in which we can follow backward the evolution of any function of time t with the help of the wave equation.

The problem of the eigenvalue representation of Hermitian operators, which was formulated in eq. (III) Sec. 6, is connected with this state of affairs. First we must investigate the region in which the operator H is defined on the manifold of the function f where $\int |f|^2 dq$ exists. We cannot obviously claim that Hf shall be meaningful everywhere, since this is not true even for the operation of multiplication by q (Note that $\int q^2 |f|^2 dq$ does not exist for all f). We can, however, insist that the manifold of functions for which Hf is defined is *everywhere* dense, i.e., for every function f , there is a function g for which Hg has a meaning, such that $\int |f - g|^2 dq$ can be made arbitrarily small. Secondly, we have to demand that H is linearly closed i.e., from $\lim_{N \rightarrow \infty} \int |f - f_N|^2 dq = 0$ and $\lim_{N \rightarrow \infty} \int |F - Hf_N|^2 dq = 0$, we should have $Hf = F$.

In thorough investigations on such Hermitian operators, von Neumann⁴ has found the remarkable result that not all such operators permit an eigenvalue representation of the form (III). For the possibility of such a representation, the reducibility of H to a diagonal form by a unitary operator U is required and, in fact, this condition is necessary and sufficient for the existence of an eigenvalue spectrum. A unitary operator U always permits an eigenvalue representation which has a form analogous to (III):⁵

$$(Uf) = \int_0^{2\pi} e^{iE} d(P_E f) \quad (8.19)$$

Here P_E and $P_E - P_{E'}$ are projection operators with the properties (I), (II) given in Sec. 6. Further the modulus of any eigenvalue of a unitary operator is 1. For unitary operators with the property (8.11) the eigenvalue representation can be given for all t in the form

$$U(t)f = \int_{-\infty}^{+\infty} e^{-i\frac{E}{\hbar}t} d(P_E f) \dots \quad (8.19')$$

For constructing a unitary operator U from H , Neumann considers, in particular, the operator

$$U = \frac{1 + iH}{1 - iH},$$

where it matters that U is meaningful everywhere. This special choice of U does not seem to be essential and we could have as well taken, e.g.,

$$U(t) = \lim_{N \rightarrow \infty} \left(1 + \frac{it}{\hbar N} H \right)^N$$

which has an obvious physical interpretation.⁶

von Neumann has shown that an exceptional operator H exists, which cannot be made unitary in this way. We shall speak only briefly about it, since it is not in any way particularly "pathological", but can be interpreted physically in a simple way. We consider a particle moving along the x -axis (one-dimensional problem), which is totally and elastically reflected at $x = 0$ by a wall. Thus only the half-plane $x > 0$ is available to it, i.e. we admit only those functions f which are defined for $0 < x < \infty$, for which the integrals $\int |f|^2 dx$ exist and which further vanish for $x = 0$; $f(0) = 0$. In this function space $v p_x = v(\hbar/t) \partial/\partial x$ is an admissible Hermitian operator (v is a constant of the dimension of velocity,

⁴ J.v. Neumann, Math. Ann. 102, 49, 370 (1929); J. reine angew. Math. 161, 208 (1929); also M.H. Stone, Proc. Nat. Acad. 15, 198, 423 (1929).

⁵ A. Wintner, Math. Z. 30, 228 (1929), as well as the book of this author, Spektraltheorie der unendlichen Matrizen, Leipzig, (1929).

⁶ Cf. also H. Weyl, Z. Physik 46, 1 (1927), and his book, Theory of Groups and Quantum Mechanics, Dover Publications Inc., New York (1950).

so that $v p_x$ has the dimension of energy). For, in the first place its domain of operation is everywhere dense; secondly it fulfils the Hermiticity condition, since $f(0) = 0$.

$$\int_0^\infty g^*(p_x f) dx - \int_0^\infty (p_x g)^* f dx = fg^*|_0^\infty = 0.$$

For this operator, however, there exists no eigenvalue representation in the space considered! This will be evident, if we consider the solutions of the equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = v p_x \psi = \frac{\hbar}{i} v \frac{\partial \psi}{\partial x} \quad \text{or} \quad \frac{\partial \psi}{\partial t} = -v \frac{\partial \psi}{\partial x}.$$

These solutions will be of the form

$$\psi(x, t) = f(x - vt). \quad (8.20)$$

$\psi(x, t)$ will be a solution in the time interval $0 < t < \tau$ only if for all t in this interval, $f = 0$ at $x = 0$. Hence $f(\xi)$ must be defined for $-vt \leq \xi < \infty$ and, further, we must have

$$f(\xi) = 0 \quad \text{for } -vt \leq \xi \leq 0. \quad (8.20')$$

i.e.

$$\psi(x, \tau) = 0 \quad \text{for } 0 \leq x \leq vt; \quad \psi(x, \tau) = f(x - vt) \quad \text{for } vt \leq x < \infty. \quad (8.20'')$$

In fact only if (8.20'') is satisfied, do we have

$$\int_0^\infty |\psi(x, \tau)|^2 dx = \int_0^\infty |\psi(x, 0)|^2 dx$$

because of (8.20). The transformation U , which takes $f(x)$ into the form $\psi(x, \tau)$ defined by (8.20), of course, preserves length but it is not unitary. This is because the manifold of $\psi(x, \tau)$ is smaller than the manifold of $f(x)$ or what is the same, the operator U^{-1} does not exist for all $f(x)$, but only for such $f(x)$ which vanish in the interval $0 \leq x \leq vt$. In the case of other functions there is no solution of the wave equation in the interval $-\tau < t < 0$. In order that the first N powers of p_x shall be well-defined in the space considered, the original function f must be modified; this modification is the more, the higher the power N is. Such an operator is obviously not admissible as a Hamiltonian. On the other hand, the operator $\partial^2/\partial x^2$ shows a normal behaviour in the space considered by us and it has the eigenfunctions $\sin kx$; so also does the operator $(\hbar/i)\partial/\partial x$ in ordinary space $-\infty < x < +\infty$. The radial momentum operator defined through $p_r f = \frac{\hbar}{ir} \frac{\partial}{\partial r} (rf)$ behaves in ordinary space in exactly the same way as the operator $(\hbar/i)\partial/\partial x$ in the half space. It is Hermitian, but its matrices cannot be diagonalised. However, the operator $p_r^2 f = -\hbar^2 \frac{1}{r} \frac{\partial^2}{\partial r^2} (rf)$ appearing in the Hamiltonian, can very well be diagonalised and has the eigenfunctions $\sin kr/r$.

CHAPTER V

Theory of Measurements

9. Determination of the Stationary States of a System Through Measurement: General Discussion of the Concept of Measurement

Before we go into a general discussion of the behaviour of systems under external perturbations, we shall give some typical examples as to how the stationary states of a system can be determined through measurement, i.e. by suitable external perturbations. Here it is essential to remember that a system, even if isolated from external influence, need not necessarily be in a stationary state or equivalently, need not necessarily possess with certainty a single value of energy. The most general state of the system would rather be given by

$$\psi = \sum_n c_n(0) e^{-\frac{2\pi i}{\hbar} E_n t} u_n(q) = \sum_n c_n(t) u_n(q), \quad (9.1)$$

with time-independent but otherwise arbitrary coefficients $c_n(0)$. (If a continuous eigenvalue spectrum is present, the sum must be replaced by an integral.) In the general case, the stationary state of the system is produced only by the measuring apparatus. We have now to investigate as to how this process can be described by the mathematical formalism of wave mechanics. The main result of the investigation is a simple statistical interpretation for the coefficients c_n .

The simplest way of investigating the state of a system (atoms or molecules), consists in placing it in an external force field, with which the different states of the system can react differently. If Q denotes the co-ordinates of the centre of mass of the molecule or atom, then any function $\psi(q, Q; t)$ can be written in the form

$$\psi(q, Q; t) = \sum_n c_n(Q; t) u_n(q, Q).$$

Here the q denote the co-ordinates of the particles of the system relative to its centre of mass. $u_n(q, Q)$ are the eigenfunctions corresponding to the energy eigenvalues $E_n(Q)$ which satisfy equations, similar to eq. (5.11):

$$-\sum_{a=1}^N \frac{\hbar^2}{2m^{(a)}} \sum_{k=1}^8 \left[\left(\frac{\partial}{\partial x_k^{(a)}} - \frac{i\epsilon^{(a)}}{\hbar c} \Phi_k(x^{(a)} + Q) \right)^2 + V^{(a)}(x^{(a)} + Q) + V(q_1, \dots, q_i) \right] u_n(q, Q) = E_n(Q) u_n.$$

$\Phi_k(q, Q)$, $V(q, Q)$ being the external potentials. In these equations, Q is treated as a parameter. If the external fields vary only slowly within the domain of definition of the system, then $\Phi_k(x^{(a)} + Q)$ and $V^{(a)}(x^{(a)} + Q)$ can be expanded in a series, which

can be terminated after a few terms and in most cases even after the first term:

$$\Phi_k(x^{(a)} + Q) = \Phi_k(Q) + \sum_{i=1}^3 \frac{\partial \Phi_k}{\partial Q_i} x_i^{(a)} + \dots,$$

$$V^{(a)}(x^{(a)} + Q) = V^{(a)}(Q) + \sum_{i=1}^3 \frac{\partial V^{(a)}}{\partial Q_i} x_i^{(a)} + \dots.$$

From the wave equation for ψ we obtain for the $c_n(Q, t)$ wave equations of the form

$$-\frac{\hbar}{i} \frac{\partial c_n}{\partial t} = -\frac{\hbar^2}{2M} \sum_{i=1}^3 \frac{\partial^2 c_n}{\partial Q_i^2} + E_n(Q) c_n, \quad (9.2)$$

disregarding the terms dropped from the above expansion. The physical significance of this wave equation is that the eigenvalue of the internal energy of the system, which is dependent on its centre of mass, appears simply as the potential energy for the motion of the centre of mass of the whole system.

The neglected terms, which will be discussed in greater detail in Sec. 11, prevent the separability of the wave equation with respect to c_n , in so far as they depend on all the c 's and their first few derivatives. Their effect is small if the motion of the system is so slow that during the time $\tau = \frac{\hbar}{E_n - E_m}$ the mean change in position $\Delta Q = \bar{Q}(t + \tau) - \bar{Q}(t)$ of the system satisfies the inequality

$$\frac{\partial E_n}{\partial Q} \Delta Q \ll E_n - E_m. \quad (9.3)$$

We should be specially careful, if in the absence of the external force field the state of the system considered is degenerate, so that certain energy differences ($E_n - E_m$) are proportional to the intensity of the external field. Even in this case the inequality (9.3) should be satisfied if the wave equation (9.2) is to be used.

Wave equations of this type form the basis for all experiments, which deal with the deflections of molecular beams in external force fields. As an example of a case in which there is degeneracy in the absence of the external field, we may mention various excited states corresponding to a total angular momentum zero of the whole system, which will be split up by an external electric field. If F is the strength of the electric field at the position Q , then in general $E_n(Q)$ will assume the form

$$E_n(Q) = -\frac{\alpha_n}{2} F^2(Q), \quad (9.4)$$

where $\alpha_n, \dots, \alpha_m, \dots$ denote the values of the electrical polarisability of the atom or molecule in the states n, m, \dots .

In the original form of the molecular beam experiment of Stern and Gerlach, in which the states corresponding to different directions of the angular momentum of the atom were split up and not the various excited states of the atoms (space quantisation), the energies of the states considered are given by

$$E_m = E_0 + \hbar \omega_L m. \quad (9.5)$$

Here m is the magnetic quantum number running from $-j$ to $+j$ (j being half integral or integral) and ω_L is a quantity proportional to the external magnetic field strength H and is equal to the Larmor frequency multiplied by a numerical factor g (the

Landé splitting factor):¹

$$\omega_L = g \frac{e \mathcal{H}}{2m_e c}. \quad (9.6)$$

The condition (9.3) states that the values of the components of \mathcal{H} along three space-fixed directions at the position of the atom can change relatively very little within the time $1/\omega_L$.

We can discuss on the basis of the wave eq. (9.2) the circumstances under which the rays (molecular or atomic) which belong to states n and m , will be spatially split up by the external force field. We imagine a cylindrical beam of diameter d to travel in the x -direction. According to (4.23), the centre of gravity of a wave packet moves along the classical path corresponding to the potential $E_n(Q)$. Now according to (4.25), (4.26) and the uncertainty relation, the spread of such a wave-packet alters because an initial minimum momentum $p_y \sim \hbar/d$ has necessarily got to be present in the transverse direction of the ray. This can also be understood as diffraction of the rays on passing through the screen of dimension d . We have now to calculate, on the one hand the deflections y_n, \dots, y_m, \dots (corresponding to the states n, m, \dots respectively), of the beams by the force field in the y -direction during the time t , and on the other, the broadening Δy of the beam in the transverse direction due to diffraction in the same time interval. In order to obtain two clearly separated rays, we must have

$$y_n - y_m \gg \Delta y. \quad (9.7)$$

Now

$$y_n = \frac{1}{2M} \frac{\partial E_n}{\partial Q_y} t^2; \quad y_m = \frac{1}{2M} \frac{\partial E_m}{\partial Q_y} t^2,$$

$$\Delta y \sim \frac{\hbar}{Md} t.$$

Hence the condition (9.7) gives

$$\frac{\partial(E_n - E_m)}{\partial Q_y} t \gg \frac{\hbar}{d}; \quad d \frac{\partial(E_n - E_m)}{\partial Q_y} t \gg \hbar. \quad (9.8)$$

The frequency $\nu_{n,m}$ corresponding to the energy difference $(E_n - E_m)$ amounts to

$$\nu_{n,m} = \frac{E_n - E_m}{\hbar}. \quad (9.9)$$

If we denote the variation of a quantity along the transverse direction of the beam by $\delta f = d \frac{\partial f}{\partial Q_y}$, then we have

$$t \delta \nu_{n,m} \gg 1. \quad (9.8')$$

Further in each case $\delta \nu_{n,m} < \nu_{n,m}$ (which is always true) so that

$$t \nu_{n,m} \gg 1. \quad (9.10)$$

We cannot determine whether the system is in the state of n or m in an arbitrarily short time, but a minimum time given by

$$t \sim \frac{1}{\nu_{n,m}} = \frac{\hbar}{E_n - E_m} \quad (9.10')$$

is required for this purpose. In the original Stern-Gerlach experiment, where the

¹ The Larmor frequency is here a circular frequency. Further it is to be remarked that even when the spin of the electron is introduced (cf. Sec. 13), the wave equation (9.2) holds under the conditions mentioned.

energy values are given by (9.5), this minimum time amounts to $1/\omega_L$. We shall see that this minimum time $1/\nu_{n,m}$ holds for all methods of determining the state of the system and not merely for the method considered here.²

The characteristic feature of the deflection experiment is that after performing the experiment the molecule or atom almost certainly lies in disjoint, possibly time-dependent regions V_n, V_m, \dots , in case the molecule or atom was certainly in the states n, m, \dots before the experiment. If, therefore, initially $c_n = 1, c_m = 0$ for $n \neq m$, the solution after the process has taken place will be

$$\psi_n(q, Q; t) = a_n(Q; t) u_n(q, Q). \quad (9.11)$$

Here the $u_n(q, Q)$ are orthogonal for fixed Q . Further if the atoms or molecules themselves have moved in a region where the external field is constant, then the Q -dependence of u drops out. As a consequence of the continuity equation, we have

$$\int a_n^* a_n dQ = 1; \quad \int a_n^* a_m dQ = 0 \quad \text{for } n \neq m \quad (9.12)$$

and due to the postulated property of the deflection experiment we have

$$a_n(Q; t) = 0 \quad \text{outside } V_n(t). \quad (9.13)$$

The linearity of the wave equation, which will be essentially used here, implies that if the system is initially in the internal state $\sum_n c_n u_n(q)$, then after the deflection experiment, the wave function $\psi(q, Q; t)$ will have the form

$$\psi(q, Q; t) = \sum_n c_n \psi_n(q, Q; t) = \sum_n c_n a_n(Q; t) u_n(q, Q). \quad (9.14)$$

According to the general principles already developed, the probability of finding the centre of mass of the system, in the interval $(Q, Q + dQ)$, whatever be the value of q , is then given by

$$W(Q) dQ = dQ \int \psi^*(q, Q; t) \psi(q, Q; t) dq = \sum_n |c_n|^2 |a_n(Q; t)|^2 dQ. \quad (9.15)$$

On account of eq. (9.13) we also find that the probability of finding the atom in the region V_n is given, in the general case, by

$$\int_{V_n} W(Q) dQ = |c_n|^2. \quad (9.16)$$

This can be related to the following statement (considered equivalent by definition): *In the general case, $|c_n|^2$ is the probability that the system is found in the state E_n .*

The justification for such a statement comes also from the result for the probability of finding the relative co-ordinates q in the interval $(q, q + dq)$ after the deflection experiment, whatever the values of Q are. We find according to eqs. (9.12) and (9.14)

$$W(q) dq = dq \int \psi^*(q, Q; t) \psi(q, Q; t) dQ = \sum_n |c_n|^2 |u_n(q)|^2 dq. \quad (9.17)$$

Here it has been assumed that the functions u_n no longer depend on Q . A similar result follows for the probability $W(p) dp$ in momentum space.

² Eq. (9.10') is not equivalent to the uncertainty relation $\Delta E \Delta t \sim \hbar$ considered earlier, since in the latter, one is concerned with the duration of time within which a particle with an uncertainty in energy given by ΔE is found at a particular place. Here, on the other hand, there is no question of any q -number (i.e. operator) entering the relation, besides E and t .

We can consider the centre of gravity of the atom as a special "measuring apparatus" (all that is required of the latter being that this brings new degrees of freedom into the system), and the energy E_n of the internal state as the quantity to be measured. Instead of the centre of gravity of the atom we could have taken any other apparatus (for which Q describes, e.g., the position of the pointer) if we stipulate that this apparatus *reacts with certainty differently* to the different states E_n , as expressed by eq. (9.13). Integration of the probability over Q (which eliminates the degrees of freedom of the apparatus) has, according to eq. (9.17), the consequence that *the phases of the amplitudes c_n that are associated with the object to be measured are no longer present in the result*. The probability of finding any quantity ξ representing the system in the interval $(\xi, \xi + d\xi)$, is then equal to the sum of these probabilities for the case, where the quantity to be measured has a particular value multiplied by suitable weight factors $|c_n|^2$:

$$W(\xi) d\xi = \sum_n |c_n|^2 W_n(\xi) d\xi \quad \left(\sum_n |c_n|^2 = 1 \right). \quad (9.17')$$

In particular this holds for conjugate quantities q and p , where $W_n(q) = |u_n(q)|^2$; $W_n(p) = |v_n(p)|^2$. In this case the totality of states is called a *quantum mechanical mixture*. In contrast to this, we have the *pure* case, for which the probability $W(\xi)$ cannot be simultaneously obtained for all quantities ξ (it is enough to consider just two conjugate quantities ξ) by addition of these probabilities in essentially possible cases. Therefore, the measurement of the internal energy E_n of the system, ignoring the degrees of freedom of the apparatus, generates in general (i.e. if not all but one of the c 's vanish) out of a pure case for which

$$W(q) dq = \left| \sum_n c_n u_n(q) \right|^2 dq; \quad W(p) dp = \left| \sum_n c_n v_n(p) \right|^2 dp.$$

a mixture for which we have

$$W(q) dq = \sum_n |c_n|^2 |u_n(q)|^2 dq; \quad W(p) dp = \sum_n |c_n|^2 |v_n(p)|^2 dp.$$

This result, which follows just from the assumptions already introduced and rests essentially on the linearity of the wave equation, is of decisive importance for the consistent interpretation of the concept of measurement in quantum mechanics. For, this result shows that we arrive at consistent results concerning the system, whatever be the way in which the division between the system to be observed (which is described by wave functions) and the measuring apparatus is made.³

It is possible to express the fact that a definite measuring apparatus will be used in the mathematical formalism of quantum mechanics directly. On the contrary, this is not possible with the stipulation that the measurement should give a definite result: for example, in our case "The centre of mass of the atom is within the region V_n after the experiment." "The energy of the atom has, therefore, the value E_n and no other." Any such statement about a physical fact made with the help of a measuring device (observer or registering apparatus) which is not counted as part of the system cannot (from the standpoint of the mathematical formalism which describes directly only probabilities) represent a particular, scientifically not pre-determined act which is to be taken into account by a reduction of the wave-packets [in our case from $\sum_n c_n u_n(q)$]

³ Cf. J.v. Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton (1955) where in Chapter VI (The Measurement Process) this question is discussed in detail.

to $u_n(q)$]. Quite similar is the situation with regard to the measurement of the position of a particle. In this case, we must substitute the function $\psi(q, t)$ for $\sum_n c_n u_n(q)$ and the position of the photon in the focal plane of the eye-piece of the γ -ray microscope for Q . As already mentioned in Sec. 1, we need not be surprised at the necessity for such a special procedure if we realise that during each measurement an interaction with the measuring apparatus ensues which is in many respects intrinsically uncontrollable. Here it is important to remember that the statement that the system necessarily possesses a definite internal energy E_n independent of its value being fixed by measurement, or, in other words, that the system is found in a definite *stationary* state, can easily give rise to contradictions, especially in cases where the old quantum theory speaks of "transition processes" between various stationary states.

All that has been said above regarding the measurement of the energy of a system is usually immediately asserted in quantum mechanics about the measurement of "any arbitrary physical quantity". We shall discuss this generalisation only later because as we have seen, such measurements require, in general, a finite, minimum time. This requires a special consideration of the fact that the quantities in question are variable in time (cf. Sec. 10). This complication is absent in the case of measurement of the energy of a (conservative) system, because the energy is constant in time. In fact, the probability that the energy of the system amounts to E_n , is given by $|c_n|^2$ and since $c_n(t) = c_n(0) \exp(-\frac{i}{\hbar} E_n t)$ we have

$$|c_n(t)|^2 = |c_n(0)|^2. \quad (9.18)$$

The probability of finding a closed system with definite energy E_n is independent of time. This is the most general statement of the law of conservation of energy. The law, verified earlier, that the expectation value \bar{E} of the energy is always constant in time is contained in the above statement as a special case, since

$$\bar{E} = \sum_n |c_n|^2 E_n. \quad (9.19)$$

It is useful to introduce here the (Hermitian) density matrix P , first defined by von Neumann,⁴ which permits us to calculate conveniently the expectation value of any quantity in a state.

Let a state be represented by the eigenfunction

$$\psi = \sum_n c_n(0) \psi_n(q, t) = \sum_n c_n(0) e^{-\frac{i}{\hbar} E_n t} u_n(q).$$

Then we can define in a matrix representation, in which $H = E$ is a diagonal matrix, the quantity

$$P_{m,n} = c_m^*(t) c_m(t). \quad (9.20)$$

Then the mean value of the energy E is given by

$$\bar{E} = \sum_n P_{n,n} E_n = \sum_n (\bar{E} P)_{n,n}$$

The mean value of q is

$$\bar{q} = \int q \psi^* \psi dq = \sum_{n,m} c_n^* c_m \int q \psi_n^* \psi_m dq = \sum_{n,m} c_n^* c_m q_{n,m},$$

or

$$\bar{q} = \sum_{n,m} q_{n,m} P_{m,n} = \sum_n (q P)_{n,n}.$$

Similarly, the mean value of an arbitrary operator F is

$$\bar{F} = \int \psi^* (F \psi) dq = \sum_{n,m} c_n^*(t) c_m(t) F_{n,m}(0) = \sum_{n,m} F_{n,m}(0) P_{m,n} = \sum_n (F P)_{n,n}.$$

⁴ J.v. Neumann, Göttinger Nachr. (1927) 245; cf. also P.A.M. Dirac, Proc. Camb. Phil. Soc. 25, 62 (1929); Proc. Camb. Phil. Soc. 26, 376 (1930) and 27, 240 (1930).

Now the *spur* (trace) of a matrix X is defined as the sum of its diagonal elements:

$$\text{Spur}(X) = \sum_n X_{n,n}. \quad (9.21)$$

The spur has the important property that the spur of the product of two matrices A and B is commutative, i.e. is independent of the order of the factors:

$$\text{Spur}(AB) = \text{Spur}(BA). \quad (9.22)$$

For,

$$\text{Spur}(AB) = \sum_{m,n} A_{n,m} B_{m,n}$$

is symmetric in A and B . If we set in (9.22) $A = S^{-1}X$, $B = S$, then we get the important relation

$$\text{Spur}(S^{-1}XS) = \text{Spur}(X). \quad (9.23)$$

which will be required later. This implies in particular that the spur of a matrix is invariant under unitary transformations.

The results obtained till now can be summarised as follows: for the most general state of a system the density matrix P determines the mean value of an operator F according to the relation

$$\bar{F} = \text{Spur}(PF) = \text{Spur}(FP). \quad (9.24)$$

In particular we can substitute for F the position co-ordinate q , the momentum co-ordinate p or the energy H of the system.

It follows from (9.23) that the expression (9.24) is invariant under a change of the representation of the matrices. For example, if q is brought to the diagonal form, we have

$$\bar{F} = \int P(q', q'') F(q'', q') dq' dq''.$$

Hence, e.g., with

$$P(q'q'') = \psi(q', t) \psi^*(q'', t); \quad F(q'q'') = F(q') \delta(q'' - q')$$

we get

$$\bar{F}(q) = \int \psi^*(q, t) F(q) \psi(q, t) dq,$$

as it should be. Further, the dependence of P on time is so chosen that, if the matrices of F are independent of time, then the mean value \bar{F} has the correct time dependence. We have, in fact, according to (9.20)

$$\frac{\hbar}{i} \dot{P}_{m,n} = P_{m,n} E_n - E_m P_{m,n}.$$

Hence in general

$$\frac{\hbar}{i} \dot{P} = -(HP - PH) \quad (9.25)$$

and this, when substituted in (9.24), gives correctly

$$\begin{aligned} \frac{\hbar}{i} \dot{\bar{F}} &= \frac{\hbar}{i} \text{Spur}(\dot{P}F) = -\text{Spur}(HPF) + \text{Spur}(PHF) \\ &= -\text{Spur}(PFH) + \text{Spur}(PHF) \\ &= -\text{Spur}(P, FH - HF) = (HF - FH) \end{aligned}$$

in agreement with eq. (8.6). We note that the sign in eq. (9.25) is the reverse of that in eq. (8.6).

Till now we have considered only *pure* cases. As is clear from (9.20) even for the most general pure case the matrix P , diagonalised by a unitary transformation, will always be equal to

$(P_n)_{mm} = \delta_{mm} \cdot \delta_{nn}$; therefore, $P_n = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix}$. That is, one of the eigenvalues of P is 1 and the others are zero. We shall see that this is also a sufficient condition for the pure case.

As von Neumann has remarked, the matrix P can be generalised such that the relations (9.24) and (9.25) hold even for mixed states. The matrix P of the most general mixture is obtained by taking linear combinations of the matrices P_1, P_2, \dots corresponding to the pure cases, according to

$$P = \sum_n p_n P_n, \quad (9.26)$$

where

$$\sum_n p_n = 1; \quad p_n \geq 0. \quad (9.26')$$

Since according to (9.26'), $\text{Spur}(\mathbf{P}_n) = 1$, it follows for all \mathbf{P}_n , that

$$\text{Spur}(\mathbf{P}) = 1. \quad (9.27)$$

This relation is, in fact, necessary according to (9.24) as we can check by taking F to be the unit matrix. We point out that \mathbf{P} is positive semi-definite, i.e. that all the non-vanishing eigenvalues of \mathbf{P} are positive.

$$\text{Negative eigenvalues for } \mathbf{P} \text{ do not exist.} \quad (9.28)$$

This statement is equivalent to the one that for all (Hermitian) operators, that are quadratic, we should have

$$\text{Spur}(\mathbf{P} \mathbf{A}^2) \geq 0. \quad (9.28')$$

In fact, if \mathbf{P} is brought to the diagonal form, it follows from (9.28') that

$$\sum_n \mathbf{P}_{n,n} \sum_k |A_{n,k}|^2 \geq 0$$

for all $A_{n,k}$; therefore, $\mathbf{P}_{n,n} \geq 0$ for all n . Conversely, from this follows the validity of (9.28') in this special representation of the matrices and because of the invariance of the trace, in any general representation. From (9.28') we see that *the sum of two positive semi-definite matrices is again positive semi-definite*. The sum of positive matrices can vanish only if each of the matrices vanishes individually. Since $p_n \mathbf{P}_n$ are positive semi-definite matrices, we see that (9.28) is, in fact, a consequence of (9.26). Conversely, the most general Hermitian matrix which satisfies the conditions (9.27) and (9.28) can be represented in the form (9.26) in which the matrices \mathbf{P}_n , corresponding to the pure cases, can be taken to be commutative. If we consider \mathbf{P} to be in the diagonal form ($\mathbf{P}_{n,n}$) we need only just set $p_n = \mathbf{P}_{n,n}$ whereas \mathbf{P}_n has the element 1 at the n^{th} position and has zeros at all the other places.

We can now bring the pure case also under the general class of matrix \mathbf{P} , which need satisfy only the conditions (9.27) and (9.28) by simply requiring

$$\mathbf{P}^2 = \mathbf{P}. \quad (9.29)$$

A Hermitian matrix satisfies this relation if and only if all its eigenvalues satisfy it, i.e. if the eigenvalues are either 0 or +1. From condition (9.27) it follows that only one of the eigenvalues is equal to 1 and all the others are 0. In the general case, $\mathbf{P} - \mathbf{P}^2$ is always a positive matrix, since the eigenvalues of \mathbf{P} are always less than or equal to 1, because of the requirement (9.27).

We next show that the composition of two ensembles to which the density matrices \mathbf{Q} and \mathbf{R} correspond according to

$$\mathbf{P} = p_1 \mathbf{Q} + p_2 \mathbf{R}; \quad (0 < p_1 < 1, 0 < p_2 < 1; p_1 + p_2 = 1)$$

can never give rise to a pure case unless $\mathbf{Q} = \mathbf{R} = \mathbf{P}$. For this purpose we construct the expression.

$$\mathbf{P}^2 = p_1^2 \mathbf{Q}^2 + p_1 p_2 (\mathbf{Q} \mathbf{R} + \mathbf{R} \mathbf{Q}) + p_2^2 \mathbf{R}^2.$$

Now

$$(\mathbf{Q} - \mathbf{R})^2 = \mathbf{Q}^2 - (\mathbf{Q} \mathbf{R} + \mathbf{R} \mathbf{Q}) + \mathbf{R}^2.$$

Therefore,

$$\mathbf{P}^2 = p_1 \mathbf{Q}^2 + p_2 \mathbf{R}^2 - p_1 p_2 (\mathbf{Q} - \mathbf{R})^2$$

(Here, since $p_1 + p_2 = 1$, we have set $p_1^2 + p_1 p_2 = p_1$, $p_2^2 + p_1 p_2 = p_2$.) Accordingly, we find that the right-hand side of the expression

$$\mathbf{P} - \mathbf{P}^2 = p_1 (\mathbf{Q} - \mathbf{Q}^2) + p_2 (\mathbf{R} - \mathbf{R}^2) + p_1 p_2 (\mathbf{Q} - \mathbf{R})^2,$$

is a sum of purely positive matrices. If \mathbf{P} is to correspond to a pure case, it follows, since $\mathbf{P}^2 = \mathbf{P}$, that all the matrices on the right-hand side vanish, and in particular

$$(\mathbf{Q} - \mathbf{R})^2 = 0.$$

The square of a Hermitian matrix can vanish only if all its elements themselves vanish. [We see this, e.g., from the relation $(\mathbf{A}^2)_{n,n} = \sum_k |A_{nk}|^2$.] Therefore,

$$\mathbf{Q} = \mathbf{R}.$$

The definition of the pure case as that ensemble for which *one* eigenvalue of the density matrix is 1 and all the others 0, is, therefore, equivalent to the statement that a pure case cannot be produced by mixing two different ensembles.

von Neumann has shown further⁵ that the quantity

$$\Sigma = \text{Spur} (\mathbf{P} \log \mathbf{P}) \quad (9.30)$$

plays, up to a factor ($1/k$) (k = Boltzmann's constant) the role of entropy associated with the density distribution \mathbf{P} . For the pure case and only for this, the trace vanishes since $\mathbf{P}_{n,n} \log \mathbf{P}_{n,n} = 0$, for the case $\mathbf{P}_{n,n} = 0$ as well as for $\mathbf{P}_{n,n} = 1$. If we assume, as usual the relations (9.27) and (9.28), we have

$$\text{Spur} (\mathbf{P} \log \mathbf{P}) = 0 \quad (9.31)$$

which is equivalent to (9.29). The distribution \mathbf{P} , which for a given mean value

$$E = \text{Spur} (\mathbf{H} \mathbf{P})$$

of the energy, makes the quantity Σ a minimum, is the canonical distribution

$$\mathbf{P} = C e^{-\mathbf{H}/\Theta}, \quad (9.32)$$

with $\Theta = kT$, where T is the temperature and C is to be evaluated from the normalisation condition (9.27). The free energy is then given by

$$e^{-F/\Theta} = \text{Spur} (e^{-\mathbf{H}/\Theta}), \quad (9.33)$$

so that eq. (9.32) can also be written as

$$\mathbf{P} = e^{(F-\mathbf{H})/\Theta}. \quad (9.32')$$

If the Hamiltonian operator \mathbf{H} can be diagonalised, then \mathbf{P} will also be diagonal and $\mathbf{P}_{n,n} = e^{-E_n/\Theta}$; therefore, according to eq. (9.33) we have

$$e^{-F/\Theta} = \sum_n e^{-E_n/\Theta}. \quad (9.33')$$

The invariance of (9.33) with respect to changes in the representation of the matrices is, however, useful in many cases.⁵

The method of measurement of the energy of the system discussed till now has the property that a repetition of measurement gives the same value for the quantity measured as in the first measurement. In other words, if the *result* of using the measuring apparatus is not known, but only the fact of its use is known (in the terminology of Sec. 1, the measured quantity is unknown after the measurement, but is determinate), the probability that the quantity measured has a certain value is the same, both before and after the measurement. We shall call such measurements as the *measurements of the first kind*. On the other hand it can also happen that the system is changed but in a controllable fashion by the measurement – even when, in the state before the measurement, the quantity measured had with certainty a definite value. In this method, the result of a repeated measurement is not the same as that of the first measurement. But still it may be that from the result of this measurement, an unambiguous conclusion can be drawn regarding the quantity being measured for the concerned system before the measurement. Such measurements, we call the *measurements of the second kind*.⁶ We have already seen in Sec. 2, that the momentum measurement of the first kind is possible only in a sufficiently long time; the momentum measurement of the second kind, however, can be made in a short time.

⁵ The treatment of the generalised quantum statistics from a quantum mechanical point of view falls outside the scope of the present work; cf. P. Jordan, Statistische Mechanik auf quantentheoretischer Grundlage, Braunschweig, (1933).*

*See also William Band, An Introduction to Quantum Statistics, D. von Nostrand Co., Inc., Princeton, New Jersey, (1955).

⁶ Cf. L. Landau and R. Peierls, Z. Physik 69, 56 (1931).

An example for energy measurement of the second kind is the effect, on an atomic system, of a collision in which the energy of the projectile is measured after the collision. If at time zero the system being measured is in a state n and if the projectile has the initial kinetic energy ϵ , then the probability, that at time t , the target system is in state m and the projectile has the kinetic energy lying between ϵ' and $\epsilon' + d\epsilon'$, is given by an expression of the form

$$W_m(\epsilon') d\epsilon' = A_{n,m} \left[\frac{1 - \cos((E_n + \epsilon - E_m - \epsilon') t/\hbar)}{E_n + \epsilon - E_m - \epsilon'} \right]^2 d\epsilon'. \quad (9.34)$$

This formula follows from the general formalism of perturbation theory [Sec. 10, in particular eq. (10.19a) and the following]. Here we assume integration over the directions of the initial and final momenta to have been already performed and the kinetic energies and the momenta to be connected by the relations $\epsilon = \frac{p^2}{2m}$, $\epsilon' = \frac{p'^2}{2m}$. The quantity inside the brackets in (9.34) is significantly different from zero at time t , only if

$$E_m - E_n - (\epsilon - \epsilon') \sim \hbar/t;$$

In case $t \gg \frac{\hbar}{|E_m - E_n|}$ as in (9.10'), the measured value of $\epsilon - \epsilon'$ will lie in the neighbourhood of one of the differences $E_l - E_n$ (for any l), if the system was in state n and in the neighbourhood of $E_l - E_m$ if it was in the state m . Under the mentioned restriction on the time, the intervals of $\epsilon - \epsilon'$, for which $W(\epsilon')$ is significantly different from zero, are clearly separated. In this way we can decide whether the system originally possessed the energy E_n or E_m .

The above experiment will be still simpler, if the system is ionised after the collision so that the energy E_m is continuous and we can write instead of eq. (9.34)

$$W(\epsilon', E') d\epsilon' dE' = A_n(E') \left[\frac{1 - \cos((E_n + \epsilon - E' - \epsilon') t/\hbar)}{E_n + \epsilon - E' - \epsilon'} \right]^2 d\epsilon' dE' \quad (9.34')$$

Then the quantities ϵ , ϵ' , and E' can be measured by known methods and we find $E' + \epsilon' - \epsilon$ to be in the neighbourhood of E_n , (the difference between them being of the order of \hbar/t) if n represents the state of the system before collision.

In the above considerations, the validity of the law of conservation of energy in collision processes is an essential factor. Further, it has been shown that within the limits of the inequality (9.10') the interaction energy of the system can be neglected for energy balance. This is again a justification to call E_n the energy of the system.

We shall now proceed to investigate the question as to what we can conclude from a collision experiment performed on a system which is in an arbitrary initial state, given by

$$\psi = \sum_n c_n u_n .$$

If the inequality (9.10') is satisfied, then the product terms on the right-hand side of (9.34) or (9.34') vanish, if they are formed for two *different* index values n and m . Therefore, we arrive by a measurement of $\epsilon - \epsilon'$ or E' and $\epsilon - \epsilon'$ directly, at a measurement of $|c_n|^2 A_{n,m}$ and $|c_n|^2 A_n(E')$. The complications arising from the factors $A_{n,m}$ and $A_n(E')$ can be removed by allowing the colliding particle to get reflected to and fro for a long time and hence to allow for repeated collisions, so that ultimately one measures its total change in energy. This change in energy will then coincide with

one of the $E_n - E_m$ (m arbitrary) in $|c_n|^2$ of the cases. In the case of ionisation and simultaneous measurement of E' , $E' - (\epsilon - \epsilon')$ will coincide with E_n , in $|c_n|^2$ cases.

We can now discuss, in a general way, the measurement of the second kind with the help of the eigenfunction ψ of the system being measured and the eigenfunction Ψ of the measuring apparatus. Let the states of the measuring apparatus which are determined correspond to a complete orthonormal set of functions U_k . In the above example we shall replace k by the energy difference $\epsilon - \epsilon'$. Since it makes no essential difference whether k is discontinuous or continuous, we shall use the notations for the discrete case and consider summations over k , even if we are actually dealing with an integral. If before measurement

$$\psi = \sum_n c_n u_n$$

is the state of the system to be measured (the u_n are orthogonal and complete), then

$$\sum_k \psi_k U_k$$

is the state of the combined system after the measurement. Further, on account of the linearity of all the Hamiltonians, ψ_k must depend linearly on the c_n :

$$\psi_k = \sum_n c_n v_k^{(n)}. \quad (9.35)$$

Here $\sum_k \int |\psi_k|^2 dQ = 1$ for all c_n ; therefore $\sum_k \int |v_k^{(n)}|^2 dq = 1$. After reading off a particular k -value on the "apparatus", the wave-packet ψ is changed (up to a constant normalisation factor) to the wave-packet ψ_k . From the measured values of k , a unique conclusion regarding c_n can be drawn if and only if for every k only one $v_k^{(n)}$ is non-zero. (The same n can, however, correspond to different k .) The states k can be broken up into separate groups such that each group belongs to a particular value of n . Hence we replace k by the double index n, m and write

$$\psi_{n,m} = c_n v_{n,m}, \quad (9.35')$$

instead of (9.35). In (9.35') for all c_n , we should have

$$\sum_{n,m} \int |\psi_{n,m}|^2 dQ = 1$$

since $\sum_n |c_n|^2 = 1$. The last relation is equivalent to

$$\int |v_{n,m}|^2 dQ = 1. \quad (9.36)$$

The probability of finding the apparatus after the measurement in the group (n, m) with a fixed n , is equal to the probability that the system was in the state n before the measurement. In agreement with this statement we also have

$$|c_n|^2 = \sum_m \int |\psi_{n,m}|^2 dQ. \quad (9.37)$$

We can expand $v_{n,m}$ also in terms of u_n :

$$v_{n,m} = \sum_l T_{l;n,m} u_l.$$

For each (n, m) , we have then, according to (9.35),

$$\sum_{l,m} |T_{l;n,m}|^2 = 1. \quad (9.38)$$

This condition is obviously much weaker than an orthogonality condition. In the measurement of the first kind T is, in particular, the unit matrix. Even for a general measurement of the second kind, the special states, for which one of the c_n is equal to 1 and the others 0, have the property that we can make a statement *with certainty* about the outcome of the measurement; i.e. the result of measurement k will fall within a particular group (n, m) with a predictable n .

We shall now return to our example of the measurement of the energy of a system through a collision process. In dealing with the excited states of the system, we shall assume that each energy difference $E_n - E_m$ is present only for one single pair of states. Then for each state k , i.e. for each $\epsilon - \epsilon'$, there is one particular E_n . The $v_{n,m}$ here are *identical with the u_m* up to constant factors, and, therefore, independent of n . In the case of ionising collisions we also consider the energy E' of the outgoing electron to be measured by the apparatus; $\epsilon - \epsilon'$ and E' together play the role of k . To each k there is an n , determined by the relation $E' - (\epsilon - \epsilon') = E_n$ while E' plays the role of m . $v_{n,m}$ is again the eigenfunction in the continuous spectrum with energy E' and is independent of n .

CHAPTER VI

Approximation Methods

10. The General Formalism of Perturbation Theory

(a) Stationary State Perturbation Theory

For many applications it is essential to have an approximate method of solving the wave equation, which can be used, if the energy matrix is not completely diagonal, but the non-diagonal elements $H_{m,n}$ are small compared to the differences between the diagonal elements:

$$|H_{m,n}| \ll |H_{m,m} - H_{n,n}| \quad (10.1)$$

Here we seek a stationary solution of the wave equation assuming that a suitable, complete orthogonal set v_1, v_2, \dots is given, for which the condition (10.1) is satisfied. For this set of functions, the wave equation for the stationary states reads:

$$\sum_n H_{m,n} c_n = c_m E. \quad (10.2)$$

The eigenfunction belonging to E is then

$$u(E) = \sum_n c_n(E) v_n,$$

since from

$$Hv_m = \sum_n v_n H_{n,m}$$

we have, on account of (10.2), the equation

$$Hu = E u.$$

If the v_n are already orthogonal and normalised and u is to be normalised, then we require that

$$\sum_n |c_n|^2 = 1. \quad (10.3)$$

From (10.2) it also follows that for different E we always have

$$\sum_n c_n^*(E) c_n(E') = 0, \quad E \neq E'. \quad (10.3')$$

If the values of E are discrete, i.e. if (10.2) can be solved only for the discrete energy values $E_1, E_2, \dots, E_n, \dots$ we can write

$$c_n(E_k) = S_{nk}.$$

instead of $c_n(E_k)$. Here on account of (10.3) and (10.3') the $S_{n,k}$ form a unitary matrix. (For continuous spectra summation over k is to be replaced by integration.)

We, therefore, introduce for an approximate solution of eq. (10.2) the assumption that the non-diagonal elements of (10.2) are small compared to the diagonal elements. To express this fact formally, we imagine that the non-diagonal elements are multiplied by a parameter ϵ . The c_n are expanded in powers of ϵ . We set

$$H_{n,n} = E_n^0 + \epsilon \Omega_{n,n}; \quad H_{m,n} = \epsilon \Omega_{m,n} \quad \text{for } n \neq m. \quad (10.4)$$

Further we seek a solution, which lies in the neighbourhood of the eigenvalue E_k^0 , i.e. a set of coefficients c_n , which is equal to $\delta_{n,k}$ in the zeroth approximation:

$$\left. \begin{aligned} E_k &= E_k^{(0)} + \epsilon E_k^{(1)} + \epsilon^2 E_k^{(2)} + \dots \\ c_{n;k} &= \delta_{n,k} + \epsilon c_{n;k}^{(1)} + \epsilon^2 c_{n;k}^{(2)} + \dots \end{aligned} \right\} \quad (10.5)$$

Substituting (10.4) and (10.5) in (10.2) and equating like powers of ϵ , we have

$$E_m^{(0)} c_{m;k}^{(1)} + \Omega_{m,k} = \delta_{m,k} E_k^{(1)} + c_{m;k}^{(1)} E_k^{(0)}, \quad (10.6a)$$

$$\left. \begin{aligned} E_m^{(0)} c_{m;k}^{(2)} + \sum_n \Omega_{m,n} c_{n;k}^{(1)} &= \delta_{m,k} E_k^{(2)} + c_{m;k}^{(1)} E_k^{(1)} + c_{m;k}^{(2)} E_k^{(0)}, \\ &\dots \end{aligned} \right\} \quad (10.6b)$$

From eq. (10.6a) follows immediately, for $m = k$,

$$E_k^{(1)} = \Omega_{k,k}. \quad (10.7a)$$

The change in the k^{th} eigenvalue is equal to the diagonal element (expectation value) of the perturbation matrix Ω in this state. For $m \neq k$, we have

$$\left. \begin{aligned} c_{m;k}^{(1)} [E_k^{(0)} - E_m^{(0)}] &= \Omega_{m,k} \\ c_{m;k}^{(1)} &= -\frac{\Omega_{m,k}}{E_m^{(0)} - E_k^{(0)}} \quad \text{for } m \neq k. \end{aligned} \right\} \quad (10.8a)$$

It is clear that the value of $c_{kk}^{(1)}$ remains undetermined. We have, however, still to consider the normalisation condition (10.3) which, when expanded in powers of ϵ , gives

$$c_{kk}^{(1)} + c_{kk}^{*(1)} = 0, \quad (10.9a)$$

$$c_{kk}^{(2)} + c_{kk}^{*(2)} + \sum_n |c_{n;k}^{(1)}|^2 = 0. \quad (10.9b)$$

From the first of these equations it follows that $c_{kk}^{(1)}$ can be an arbitrary, purely imaginary number. This indeterminacy corresponds to the circumstance, that in the solution of (10.2), the phase constants remain always arbitrary; if $c_{n;k}$ is a solution, then so is

$$c'_{n;k} = c_{n;k} e^{i\delta_k}$$

with arbitrary δ_k . Further we can set

$$\delta_k = \epsilon \delta_k^{(1)} + \epsilon^2 \delta_k^{(2)} + \dots$$

where $\delta_k^{(1)}, \delta_k^{(2)}, \dots$ are arbitrary, without contradicting (10.5).

We now proceed to a discussion of the second approximation. It follows from (10.6b) for $m = k$, remembering (10.7a), that

$$E_k^{(2)} = \sum_n \Omega_{k,n} c_{n;k}^{(1)} = - \sum_n \frac{\Omega_{k,n} \Omega_{n,k}}{E_n^0 - E_k^0} = - \sum_n \frac{|\Omega_{k,n}|^2}{E_n^0 - E_k^0}. \quad (10.7b)$$

The prime on the summation sign signifies that in the summation the value $m = k$ is to be left out. For the lowest state k this perturbation of the eigenvalue is always negative. For $m \neq k$ it follows from (10.6b) that

$$\begin{aligned} c_{m;k}^{(2)} [E_m^0 - E_k^0] &= -(\Omega_{m,m} - \Omega_{k,k}) c_{m;k}^{(1)} - \sum'_{n \neq m} \Omega_{m,n} c_{n;k}^{(1)}, \\ c_{m;k}^{(2)} &= \frac{(\Omega_{m,m} - \Omega_{k,k}) \Omega_{m,k}}{(E_m^0 - E_k^0)^2} + \sum'_{\substack{n \neq m}} \frac{\Omega_{m,n} \Omega_{n,k}}{(E_m^0 - E_k^0)(E_n^0 - E_k^0)} \quad \text{for } m \neq k. \end{aligned} \quad (10.8b)$$

$c_{kk}^{(2)}$ does not appear in these equations and need satisfy only the equation (10.9b). We may formulate these results still more transparently, if we introduce the Hermitian matrix T with the elements

$$T_{k,k} = 0; \quad T_{m,k} = i \frac{\Omega_{m,k}}{E_m^0 - E_k^0} \quad \text{for } m \neq k. \quad (10.10)$$

Then

$$c_{m;k}^{(1)} = i T_{m,k} \quad \text{for } m \neq k, \quad (10.8a')$$

$$E_k^{(2)} = +i \sum_n \Omega_{k,n} T_{n,k} \quad (10.7b')$$

$$c_{m;k}^{(2)} = -i \frac{\Omega_{m,m} - \Omega_{k,k}}{E_m^0 - E_k^0} T_{m,k} - (T^2)_{m,k} \quad \text{for } m \neq k. \quad (10.8b')$$

We recall here the lemma that an infinitesimal unitary transformation S can always be represented in terms of a Hermitian matrix T in the form

$$S = 1 + \epsilon i T. \quad (10.11)$$

Then the condition $S \tilde{S} = \tilde{S} S = 1$ is equivalent, on neglecting higher powers of ϵ , to

$$T = \tilde{T}, \quad (10.11')$$

i.e. to the hermiticity of T (cf. for this, Sec. 8). To this corresponds the fact that on account of (10.9a) and (10.8a'), $c_{kn}^{(1)}$ is equal to a Hermitian matrix up to a factor i .

From the form (10.8a) of $c_{m;k}^{(1)}$, and remembering (10.4), it can be seen that the condition (10.1) is decisively important for employing our expansion method. If this condition is violated, then the smallness of ϵ is not enough to justify the method, since $\epsilon c_{m;k}^{(1)}$ will then be of the order of magnitude 1. This is especially the case, when the system considered is degenerate, i.e. if the eigenvalue E_n^0 is repeated. In this case it is necessary to consider the corresponding finite-dimensional sub-space separately and the eigenvalue problem

$$\sum_{n=1}^g H_{m,n} c_n = c_m E; \quad m = 1, 2, \dots, g \quad (10.12)$$

is to be solved in the relevant g -dimensional sub-space in which $E_n^0 - E_m^0$ is of the same order of magnitude as $\Omega_{m,n}$. This is a purely algebraic problem and is always soluble, e.g. the g new eigenvalues, E , are determined by the condition that the determinant (the so-called "secular determinant") vanishes:

$$\begin{vmatrix} H_{11} - E, H_{12}, \dots, H_{1g} \\ H_{21}, H_{22} - E, \dots, H_{2g} \\ \vdots & & \vdots \\ H_{g1}, \dots, H_{gg} - E \end{vmatrix} = 0. \quad (10.13)$$

This equation of degree g in E always possesses g real roots in the case of a Hermitian matrix $H_{n,m}$. Now carry out the transformation $\bar{v}_m = \sum_n c_{n,m} v_n$ with $c_{n,m}$ which are determined by (10.12) with $E = E_m$. These $c_{n,m}$ can be considered as an adaptation of the basis v_n to the perturbation Ω . Then the original perturbation method can be employed, since by this transformation, the $\Omega_{m,k}$ vanish, if m and k both lie in the same sub-space. Eqs. (10.8a) and (10.7b) can again be used, if the conditions $m \neq k$ and $n \neq k$ are generalised so as to mean that the states m and k as well as n and k belong to entirely different unperturbed energies (i.e. they lie in different finite sub-spaces considered above) so that for these pairs of states the inequality (10.1) is again fulfilled.

We briefly mention here the way in which the above perturbation calculation is carried out if the energy eigenvalues lie in the continuous spectrum. Instead of the index n , we have the continuous parameter n , so that

$$u(k) = \int c(n, k) v(n) dn,$$

$$Hv(m) = \int v(n) H(n, m) dn,$$

$$\int H(m, n) c(n, k) dn = c(m, k) E(k).$$

The way in which E depends on k is still arbitrary and can be chosen according to the requirements of the calculation. In the place of (10.4) we set

$$H(m, n) = E_m^0 \delta(m - n) + \epsilon \Omega(m, n),$$

where the δ -function defined by (6.25) replaces the Kronecker symbol δ_{mn} appearing in the discrete case.

$$c(n, k) = \delta(n - k) + \epsilon c^{(1)}(n, k) + \epsilon^2 c^{(2)}(n, k) + \dots$$

Similarly (10.6a) takes the form

$$[E^0(m) - E^0(k)] c^{(1)}(m, k) = - [\Omega(m, k) - E^{(1)}(k) \delta(m - k)].$$

The appearance of the δ -function does not permit immediately to specialise to the case $m = k$ and $E^{(1)}(k)$ remains arbitrary. If $\Omega(m, k)$ possesses no singularity for $m = k$ or more precisely, if $\int_{m-\epsilon}^{m+\epsilon} \Omega(m, k) dk$ vanishes as $\epsilon \rightarrow 0$, it is convenient to set $E^{(1)}(k) = 0$. Therefore, we write

$$\Omega(m, k) - E^{(1)}(k) \delta(m - k) = \Omega'(m, k)$$

and insist that

$$\lim_{\epsilon \rightarrow 0} \int_{m-\epsilon}^{m+\epsilon} \Omega'(m, k) dk = 0.$$

Then

$$c^{(1)}(m, k) = - \frac{\Omega'(m, k)}{E^0(m) - E^0(k)}$$

will be singular for $m = k$. A closer investigation shows that it is always permissible to substitute, for an integral of the type

$$\int_{k-a}^{k+a} f(m) c^{(1)}(m, k) dm,$$

where $f(m)$ is continuous, but otherwise arbitrary, its principal value. The principal value is defined by

$$H \int_{k-a}^{k+a} = \lim_{\epsilon \rightarrow 0} \left[\int_{k-a}^{k-\epsilon} + \int_{k+\epsilon}^{k+a} \right].$$

or by

$$H \int F(m, k) dm = \frac{1}{2} \int_{k-a}^{k+a} [F(m, k) + F(2k - m, k)] dm,$$

where the integrand is regular. This holds for the calculation of the quantity $u(k)$ from $v(n)$ through the $c(n, k)$ as also for the calculation of $E^{(2)}(k)$ and $c^{(2)}(m, k)$.

(b) Time-dependent Perturbation Theory

We shall now consider time-dependent perturbations.* Here we search for a given initial state ($t = 0$), a solution of the equation

$$-\frac{\hbar}{i} \dot{c}_m = \sum_n H_{m,n} c_n, \quad (10.14)$$

where the matrix elements of the Hamiltonian H constructed from a time-independent orthogonal system are given by

$$H_{m,n} = E_n \delta_{m,n} + \varepsilon \Omega_{m,n}(t). \quad (10.15)$$

Here the dependence of the perturbation matrix element $\Omega_{m,n}$ on time is assumed to be known. The unperturbed solution then reads

$$c_n^{(0)}(t) = c_n^{(0)}(0) e^{-\frac{i}{\hbar} E_n^0 t},$$

and we seek a perturbed solution

$$c_n(t) = c_n^{(0)}(t) + \varepsilon c_n^{(1)}(t) + \varepsilon^2 c_n^{(2)}(t) + \dots$$

with the given initial values of $c_n(0) = c_n^{(0)}(0)$, so that $c_n^{(1)}(0) = c_n^{(2)}(0) = \dots = 0$. It is convenient to separate the factor $e^{-\frac{i}{\hbar} E_n^0 t}$ in c_n :

$$c_n(t) = a_n(t) e^{-\frac{i}{\hbar} E_n^0 t} \quad (10.16)$$

and to set

$$\Omega'_{m,n}(t) = \Omega_{m,n}(t) e^{\frac{i}{\hbar} (E_m^0 - E_n^0)t}. \quad (10.17)$$

Then we have

$$-\frac{\hbar}{i} \dot{a}_m = \varepsilon \sum_n \Omega'_{m,n}(t) a_n(t). \quad (10.15')$$

Therefore, with

$$a_n(t) = a_n^{(0)}(t) + \varepsilon a_n^{(1)}(t) + \dots; \quad (a_n^{(0)}(t) = a_n^{(0)}(0) = \text{const}),$$

we get

$$-\frac{\hbar}{i} \dot{a}_m^{(1)} = \sum_n \Omega'_{m,n}(t) a_n^{(0)}(0),$$

$$-\frac{\hbar}{i} \dot{a}_m^{(2)} = \sum_n \Omega'_{m,n}(t) a_n^{(1)}(t)$$

These equations can be immediately integrated to give

$$a_m^{(1)}(t) = -\frac{i}{\hbar} \sum_n a_n^{(0)}(0) \int_0^t \Omega'_{m,n}(t') dt, \quad (10.18a)$$

$$\left. \begin{aligned} a_m^{(2)}(t) &= -\frac{i}{\hbar} \sum_l \int_0^t \Omega'_{m,l}(t') a_l^{(1)}(t') dt \\ &= -\frac{1}{\hbar^2} \sum_n a_n^{(0)}(0) \sum_l \int_0^t \int_0^t \Omega'_{m,l}(\tau) d\tau \int_0^t \Omega'_{l,n}(\tau') d\tau' \end{aligned} \right\} \quad (10.18b)$$

* This method is usually referred to as the method of variation of constants and was introduced by Dirac - P.A.M. Dirac, Proc. Roy. Soc. (London). A 112, 661 (1926), A 114, 243 (1927).

An important special case occurs when the $\Omega_{m,n}$ are independent of time so that from eq. (10.17), we have

$$\Omega'_{m,n}(t) = \Omega_{m,n}(0) e^{\frac{i}{\hbar} (E_m^0 - E_n^0)t}.$$

Then, according to (10.18a,b), for time-independent case, we have

$$a_m^{(1)}(t) = - \sum_n a_n^0(0) \Omega_{m,n}(0) \frac{e^{\frac{i}{\hbar} (E_m^0 - E_n^0)t} - 1}{E_m^0 - E_n^0}, \quad (10.19a)$$

$$a_m^{(2)}(t) = \pm \sum_n a_n^0(0) \sum_l \Omega_{m,l}(0) \Omega_{l,n}(0) \left[\frac{e^{\frac{i}{\hbar} (E_m^0 - E_n^0)t} - 1}{(E_m^0 - E_n^0)(E_l^0 - E_n^0)} - \frac{e^{\frac{i}{\hbar} (E_m^0 - E_l^0)t} - 1}{(E_m^0 - E_l^0)(E_l^0 - E_n^0)} \right]. \quad (10.19b)$$

If in (10.19a) $E_m^0 = E_n^0$, then the n^{th} term takes the form

$$a_n^0(0) \Omega_{m,n}(0) \frac{i}{\hbar} t$$

and this can be used so long as $|\epsilon| |\Omega_{m,n}(0)| t / \hbar \ll 1$. If we wish to have, in this case, which corresponds to the vanishing of the denominator (resonance denominator) a solution which is valid for large values of t , then we have to modify the perturbation method, as is done in the case of the stationary solutions of degenerate or nearly degenerate systems.

A case occurring often is one in which a single discrete energy value of the unperturbed system is in a region in which the system also possesses a continuous eigenvalue spectrum (pre-dissociation; Auger effect). Then in (10.19a) the index n is allowed to be discrete while m takes continuous values and $\Omega_{m,n}$ is the matrix element normalised with respect to the parameter m . (The case, when there are many parameters, is quite similar.) We can then ask for the probability that the system has made a transition at time t to any one of the states, m for which

$$E_n^0 - \Delta E < E^0(m) < E_n^0 + \Delta E,$$

if at time $t = 0$ it was certainly in the discrete state n ($a_n^0(0) = 1$; $a_n^0(m; 0) = 0$). This transition probability $W(t)$ is given by

$$W(t) = \int |a^{(1)}(m, t)|^2 dm = \int dm |\Omega_{m,n}(0)|^2 \frac{4 \sin^2 \left[\frac{(E^0(m) - E_n^0)t}{2\hbar} \right]}{[(E^0(m) - E_n^0)^2]}.$$

Here the integral extends over the region of the m -space which corresponds to the energy interval $(E_n^0 - \Delta E, E_n^0 + \Delta E)$. If

$$\frac{\Delta E \cdot t}{\hbar} \gg 1, \quad (10.20)$$

then, on introducing

$$\frac{[E^0(m) - E_n^0]t}{2\hbar} = x$$

as the new variable of integration, $|\Omega_{m,n}(0)|^2$ can be taken outside the integral sign, to a fair degree of approximation and the resulting integral can then extend from $-\infty$ to $+\infty$. In this way we get

$$W(t) = |\Omega_{m,n}(0)|^2 \frac{t}{2\hbar} \frac{dm}{dE^0(m)} 4 \int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2} dx.$$

Since the above integral has the value π ,

$$W(t) = \frac{2\pi t}{\hbar} |\Omega_{m,n}(0)|^2 \frac{dm}{dE^0(m)}. \quad (10.21)$$

The term $dm/dE^0(m)$ effects the transition from the normalisation of the matrix element $\Omega_{m,n}$ with respect to m to normalisation with respect to $E^0(m)$. If more than one parameter, say m_1, m_2, m_3 , are present, then we must consider the volume element of the energy shell $E_n^0 - \Delta E < E(m_1, m_2, m_3) < E_n^0 + \Delta E$ in the m -space. If this is given by $\int dm_1 dm_2 dm_3 = \omega(E_n^0) 2\Delta E$, ($E_n^0 - \Delta E < E(m_1, m_2, m_3) < E_n^0 + \Delta E$), then the factor $\omega(E_n^0)$ appears in the place of $dm/dE^0(m)$, in (10.21).

The statements regarding the transition probability in a collision process which are made use of in eq. (9.34) are always contained in the general formula (10.19a). In this case the unperturbed system consists of two independent sub-systems such that the total energy of the system in eq. (10.19a) is equal to the sum of the energies of the two sub-systems. So E_n^0 is to be replaced by $E_n + \epsilon$ and $E^0(m)$ by $E_m + \epsilon'$, if E_n and E_m relate to the atom (target) only and ϵ and ϵ' to the projectile.

Instead of starting from a discrete state we could have started from a continuum state as the initial state. We get from (10.19) again a formula of the type (10.21):

$$W_n(t) = \frac{2\pi t}{\hbar} |\Omega_{n,m}(0)|^2 \frac{dm}{dE^0(m)} P(m_0). \quad (10.21')$$

$P(m)dm$ denotes the density of the systems in m -space. The interpretation is: We consider a large number of systems, the fraction $P(m)dm$ of which possesses a m -value between m and $m + dm$. Also m_0 denotes that value of m for which $E(m_0) = E_n^0$. Here, however, we have to average over the phases of $a_m^0(0)$. As $|a_m^0(0)|^2$ is independent of m in the interval considered, it can be taken outside the integral sign and set equal to $P(m_0)$. On account of the Hermiticity of Ω we have $|\Omega_{m,n}(0)|^2 = |\Omega_{n,m}(0)|^2$ and this determines, according to (10.21) and (10.21'), a relation between the frequency of a process and that of the inverse process, which is of importance for considerations of thermal equilibrium.¹

The "transition processes" considered here represented already in the old quantum theory a case where the causal description of nature could not be implemented, especially when for the same initial state more than one different transitions are possible, from which a choice is apparently made purely at random. Quantum mechanics does not, strictly speaking, "know" the concept of the (discontinuous) "process" since all the *temporal* changes of the state take place continuously. Only observation (measurement) determines to which state the system has actually gone to and the discontinuity imposed by the finiteness of the quantum of action lies exclusively in the reduction of the wave-packet (which is symbolic and describes the system only statistically) necessary for the separation of the system observed from the means of observation.

¹ This relation holds only in the first approximation of the perturbation calculation, while thermal equilibrium can be derived from more general assumptions also. Cf. E.C.G. Stueckelberg, Helv. Phys. Acta 25, 577 (1952); M. Inagaki, G. Wanders and C. Piron, Helv. Phys. Acta 27, 71 (1954).

11. Adiabatic and Sudden Perturbations

(a) Adiabatic and Sudden Perturbations of a System

A special case of the influence of an external agency on a system arises when the influence can be described by an alteration of parameters (external field strengths, location of walls, etc.). In the old quantum theory there already existed the well-known *adiabatic principle* of Ehrenfest¹ which says that a system which initially finds itself in a definite stationary quantum state remains in this state, provided the alteration of the parameters of the system takes place sufficiently slowly. The fact that such a law holds in quantum mechanics was first formulated and proved by Born.²

We, therefore, take the Hamiltonian H to be dependent on a parameter a and consider the eigenvalue problem to be solved for all the values of a which are of interest in terms of the eigenvalues $E_n(a)$ and the eigenfunctions $u_n(a)$. These, therefore, satisfy the equations

$$H(\mathbf{p}, \mathbf{q}, a) u_n(a) = E_n(a) u_n(a), \quad (11.1)$$

identically for every a . By differentiating eq. (11.1) with respect to a , we obtain

$$\left(\frac{\partial H}{\partial a} \right)_{\mathbf{p}, \mathbf{q}} u_n(a) + H \frac{\partial u_n}{\partial a} = \frac{\partial E_n}{\partial a} u_n(a) + E_n \frac{\partial u_n}{\partial a}. \quad (11.1')$$

We now set

$$\frac{\hbar}{i} \int u_m^* \frac{\partial u_n}{\partial a} dq = k_{mn}. \quad (11.2)$$

where k_{mn} is Hermitian due to the orthogonality and normalisation of u . Further, because of the Hermicity of H

$$\frac{\hbar}{i} \int u_m^* \left(H \frac{\partial u_n}{\partial a} \right) dq = \frac{\hbar}{i} \int (H u_m)^* \frac{\partial u_n}{\partial a} dq = E_m k_{mn}.$$

Therefore, on multiplying (11.1') by $(\hbar/i) u_m^*$ and integrating over q -space, we obtain

$$\frac{\hbar}{i} \left(\frac{\partial H}{\partial a} - \frac{\partial E}{\partial a} \right) = \mathbf{k} \mathbf{E} - \mathbf{E} \mathbf{k}. \quad (11.3)$$

This equation is to be understood as a matrix equation in which \mathbf{E} is a diagonal matrix. In this case the diagonal elements of the right-hand side vanish identically so that

$$\left(\frac{\partial H}{\partial a} \right)_{nn} = \frac{\partial E_n}{\partial a} \quad (11.3')^*$$

¹ P. Ehrenfest, Ann. d. Phys. 51, 327 (1916). Later Bohr discussed in particular the question of the applicability of classical mechanics for adiabatic (indefinitely slow) processes. This aspect of the problem is no longer of interest, since classical mechanics fails to describe the quantum states.

² M. Born, Z. Physik 40, 167 (1926). Later papers on this standpoint are: E. Fermi and F. Persico, Rend. Lincei (6) 4, 452 (1926); M. Born and V. Fock, Z. Physik 51, 165 (1928); P. Güttinger, Z. Physik 73, 169 (1931).

*This result is sometimes referred to as Hellmann-Feynman theorem! See, e.g., E. Merzbacher, Quantum Mechanics, Second edn., John Wiley and Sons, (1970), p. 442.

and

$$k_{mn} = \frac{1}{E_n - E_m} \frac{\hbar}{i} \left(\frac{\partial H}{\partial a} \right)_{mn} \quad \text{for } m \neq n (E_m \neq E_n). \quad (11.3'')$$

If a varies with time, we seek a solution of the equation

$$-\frac{\hbar}{i} \dot{\psi} = H(\mathbf{p}, \mathbf{q}, a(t)) \psi$$

and inserting

$$\psi = \sum_n c_n(t) u_n(a)$$

into

$$-\frac{\hbar}{i} \int u_m^* \dot{\psi} dq = \int u_m^* H \psi dq = E_m(a) \int u_m^* \psi dq,$$

we have

$$-\frac{\hbar}{i} \dot{c}_m + \dot{a} \sum_n k_{mn} c_n = E_m(a) c_m. \quad (11.4)$$

A deeper discussion of the equation shows³ that

$$c_m(T) - c_m(0) = \dot{a} F, \quad (11.5)$$

in which for fixed and finite $\dot{a}T = a(T) - a(0)$ and in the limit, $T \rightarrow \infty$, and hence for $\dot{a} \rightarrow 0$, F remains finite. Here it is assumed that during the process none of the differences $E_n - E_m$ passes through zero. This exceptional case has been discussed in particular by Born and Fock and in this discussion a reasoning due to von Laue⁴ plays an important role. Even in this case, for fixed $a(T) - a(0)$, it holds that

$$\lim_{\dot{a} \rightarrow 0} (c_m(T) - c_m(0)) = 0. \quad (11.5')$$

From (11.5) it follows that $|c_m(T) - c_m(0)|^2$ will be of the order of magnitude \dot{a}^2 for small \dot{a} .

In particular it holds in the case $c_m(0) = 0$ that $|c_m(T)|^2 \sim \dot{a}^2$. The frequency of the transition from one stationary state to another, which is caused by the change in the parameter a (jerk effect) is, therefore, proportional to \dot{a}^2 .

A more general case than those considered till now consists in adding new degrees of freedom to the system through the parameter a . If, e.g., an atom moves through a spatially varying force field, we can, in the first approximation, consider the nucleus to be infinitely heavy and the eigenvalue problem to be solved for each position Q of the nucleus:

$$H_0(q, Q) u_n(q, Q) = E_n(Q) u_n. \quad (11.6)$$

Here Q is not to be considered as a function of time, but there are new degrees of freedom which correspond to different values of Q .

Now we have to solve the equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \left(-\frac{\hbar^2}{2M} \sum_k \frac{\partial^2}{\partial Q_k^2} + H_0 \right) \psi(q, Q). \quad (11.7)$$

where H_0 acts only on q and not on Q . The same problem occurs also for molecules if the nuclei are considered fixed ($M = \infty$) in the first approximation and only in the second approximation is their motion (vibration and rotation) considered. If we set

$$\psi(q, Q) = \sum_n \varphi_n(Q, t) u_n(q, Q), \quad (11.8)$$

³ Cf. P. Ehrenfest, loc. cit.

⁴ M.v. Laue, Ann. d. Phys. 76, 619 (1925).

then, according to relation (11.6), we get

$$\begin{aligned} -\frac{\hbar}{i} \sum_n \frac{\partial \varphi_n}{\partial t} u_n(q, Q) &= \sum_n \left\{ \left[-\frac{\hbar^2}{2M} \sum_k \frac{\partial^2 \varphi_n}{\partial Q_k^2} + E_n(Q) \varphi_n \right] u_n(q, Q) - \right. \\ &\quad \left. - \frac{\hbar^2}{2M} \left[2 \sum_k \frac{\partial \varphi_n}{\partial Q_k} \frac{\partial u_n}{\partial Q_k} + \varphi_n \sum_k \frac{\partial^2 u_n}{\partial Q_k^2} \right] \right\}. \end{aligned}$$

If we introduce the Hermitian matrices

$$A_{mn}^{(h)} = \frac{1}{M} \frac{\hbar}{i} \int u_m^* \frac{\partial u_n}{\partial Q_k} dq; \quad B_{mn} = -\frac{\hbar^2}{2M} \int u_m^* \sum_k \frac{\partial^2 u_n}{\partial Q_k^2} dq, \quad (11.9)$$

then, defining the perturbation operator Ω by

$$\Omega \varphi_m = \sum_n \left(\sum_k A_{mn}^{(h)} \frac{\hbar}{i} \frac{\partial \varphi_n}{\partial Q_k} + B_{mn} \varphi_n \right), \quad (11.10)$$

we have

$$-\frac{\hbar}{i} \frac{\partial \varphi_m}{\partial t} = -\frac{\hbar^3}{2M} \sum_k \frac{\partial^2 \varphi_m}{\partial Q_k^2} + E_m(Q) \varphi_m + \Omega \varphi_m. \quad (11.11)$$

In particular we get the stationary solutions

$$\varphi_m(Q, t) = v_m(Q) e^{-\frac{i}{\hbar} Et}$$

for which

$$-\frac{\hbar^3}{2M} \sum_k \frac{\partial^2 v_m}{\partial Q_k^2} + E_m(Q) v_m + \Omega v_m = E v_m. \quad (11.11')$$

Here E can have either a discrete or a continuous spectrum or a mixture of both.

Under certain circumstances the perturbation operator Ω can be considered small and the usual perturbation method can be applied to (11.11) or (11.11'). The zeroth approximation will be the solution of the equations

$$-\frac{\hbar}{i} \frac{\partial \varphi_m^0}{\partial t} = -\frac{\hbar^3}{2M} \sum_k \frac{\partial^2 \varphi_m^0}{\partial Q_k^2} + E_m(Q) \varphi_m^0 \quad (11.12)$$

and

$$-\frac{\hbar^2}{2M} \sum_k \frac{\partial^2 v_m^0}{\partial Q_k^2} + E_m(Q) v_m^0 = E v_m^0. \quad (11.12')$$

This perturbation method can be applied to the treatment of vibration of the nuclei in a molecule, the passage of an atom through an external force field and to other problems.⁵ We can call this approximation as the *adiabatic approximation*, because in this approximation the system always remains in the same internal state m and also because the approximation is better, the smaller the variation of the wave packet v_m^0 in the course of a vibration period $\hbar/[E_m(Q) - E_n(Q)]$. In the case of the discrete spectrum of eq. (11.12') the approximation is better, the smaller the energy differences $E' - E''$ of the spectrum are, compared to the differences $E_m - E_n$ of the spectrum of the internal energy. (The vibration frequency of the nucleus is small compared to the electron frequency.)

The equations (11.11) and (11.11') form also the bases for studying the passage of atomic beams through magnetic fields whose directions vary in space.⁶ In this case, it suffices to consider in eq. (11.10) a finite number of states corresponding to the quantisation of the directions. From the explicit existence of stationary solutions according to (11.10') it follows further that even when the jerk-effect is considered (only in so far as the external fields are not varying with time), the sum of the internal energy and the translational energy remains constant.

Of particular interest, besides the limiting case of adiabatic processes, is the case

⁵ Cf. M. Born and J.R. Oppenheimer, Ann. d. Phys. **84**, 457 (1927). For general methods cf., further J. Frenkel, Phys. Z. Sowjet, **1**, 99 (1932). Also L. Landau, Phys. Z. Sowjet, **1**, 88 and **2**, 46, (1932).

⁶ Cf. for this C.G. Darwin, Proc. Roy. Soc. Lond., Ser. A **117**, 258 (1927), especially Sec. 10.

of a "sudden" change of the parameter a . Here the meaning of the word "sudden" is made precise by saying that the relative change of a , during the period $\frac{1}{\nu_{mn}} = \frac{\hbar}{E_n - E_m}$ in question must be large:

$$|\dot{a}| \gg a \left| \frac{E_n - E_m}{\hbar} \right| \quad (11.13)$$

Then in the first approximation and for a finite change of the parameter a one can set equal to zero all quantities proportional to t in the equation

$$-\frac{\hbar}{i} \frac{\partial}{\partial t} \sum_n c_n u_n = \sum_m E_m c_m u_m$$

after integration over the duration of the change:

$$-\frac{\hbar}{i} \sum_n c_n u_n(a) \Big|_0^t = \int_0^t \sum_m E_m c_m u_m dt.$$

Since this holds for the right-hand side of the equation, it must also be true for the left-hand side, so that we have

$$\sum_n c_n u_n = \sum_n c_n(0) u_n(0),$$

i.e. continuity of the function $\psi = \sum c_n u_n$ during the sudden change of the parameter a , where

$$c_m(t) = \sum_n S_{mn} c_n(0) \quad (11.14)$$

with

$$S_{mn} = \int u_m^*(a) u_n(0) dq.$$

If H_{mn} are the matrix elements of the new Hamiltonian (corresponding to the parameter a) taken between the set of eigenfunctions of $H(0)$, i.e.

$$H_{mn} = \int u_m^*(0) H(a) u_n(0) dq, \quad (11.15)$$

it holds that

$$E(a) = S H S^{-1}. \quad (11.16)$$

That is, S brings the matrix of the new Hamiltonian to the diagonal form.

(b) The Most General Statement on Probability in Quantum Mechanics

We can now discuss the general probability statements of quantum mechanics in which we consider for the sake of simplicity of notation only quantities with discrete eigenvalues. This can be formulated in the following way: We ask for the probability $W(F_n; G_m)$ that at a definite instant of time t_1 , a certain quantity F assumes the particular value F_n if earlier, at time $t_0 = t_1 - \tau$, another quantity G had assumed the value G_m . If S is the transformation matrix which takes one from a matrix representation in which $F(t_0)$ is brought to the diagonal form, to the representation in which $G(t_1)$ is also diagonalised, then the required probability is given by

$$W(F_n; G_m) = |S_{nm}|^2. \quad (11.17)$$

(It is, therefore, symmetric with respect to the quantities F and G .) This statement is

contained in our earlier statements, if (a) one or both of the quantities F and G are position- or momentum-variables (possibly at different instants of time, cf., Sec. 3, 4, 5 and 9) or (b) one or both of the quantities F and G commute with the Hamiltonian of the system, and, therefore, must be constant in time.

If neither of these two possibilities is pertinent, one can have recourse to a trick which forms a third possibility for the measurement of the quantities, and which rests on the above-mentioned "sudden" change in the Hamiltonian. The statement (11.17) is also valid if (c) it is possible by changing a parameter, e.g., switching off the external field, to make the quantity concerned "suddenly" constant in time (in the sense explained above), i.e. to change the Hamiltonian such that the quantity concerned can commute with it ("stop assumption"). In this case we can first make the quantity F constant in time at time t_0 and measure it, and after a time τ do the same with the quantity G . The validity of (11.17) follows for this case from the result proved above regarding the sudden change together with the results proved in Sec. 9.

It must, however, be stressed that this sudden stopping of a quantity is possible only in very restricted circumstances. For example, it is impossible to 'switch off' the charge of the proton suddenly in order to make the momentum of the electron in a H-atom suddenly constant in time. In this case, one succeeds, of course, in the determination of the eigenfunction $\varphi_n(p)$ in the momentum space by a measurement of the second kind, the direct repetition of which would give a different result [possibility (a) above]. It has not been proved in general that every quantity can be measured in an arbitrarily short time, even when one permits measurements of the second kind.

On account of this, we prefer, in contrast to the dogmatic basis of transformation theory, not to introduce the statement (11.17) as an axiom. For, ultimately, the measurement of a quantity, in so far as it is feasible, can be reduced to the measurement of position on an apparatus. An apparatus measures a quantity F , when in the decomposition of the ψ -function into the normalised orthogonal functions u_n of the corresponding operator F , the apparatus shows with certainty the "position of the needle" Q_1, Q_2, \dots, Q_n , if ψ was equal to u_1, u_2, \dots, u_n respectively before the measurement. In the general case the probability for the position of the needle Q_n is to be defined as the probability that the quantity F had the value F_n before the measurement. From the significance of the ψ -function of the apparatus and the linearity of all Hamiltonians it then follows that this probability is equal to the square of the absolute value of the expansion coefficient c_n of the ψ -function of the system to be measured (before measurement) in terms of the function u_n , ($c_n = \int u_n^* \psi_n dq$). Suppose later on a new measurement of the other quantity G is made with a new apparatus which can be described by new "needle positions" Q_1, Q_2, \dots . Then the probability that the new quantity G has a certain value G_m , if earlier the quantity F certainly had the value F_n , is defined as identical with the probability that the new apparatus shows the needle position Q_m , if it is known, that the first apparatus, which must allow a unique conclusion to be reached on the state after the measurement (which in measurements of the second kind is not the same as the state before the measurement) has, with certainty, the needle position Q_n . The latter statement is, however, equivalent to saying that the state of the system being measured will be

described by the eigenfunction u_n after the first measurement.⁷ If $v_1, v_2, \dots, v_m, \dots$ are the eigenfunctions of the quantity G , then the probability $W(F_1, G_m)$ is in fact equal to $|S_{nm}|^2$, where $S_{nm} = \int u_n v_m^* dq$.

Thus all the statements regarding the arbitrary quantities F and G appear to reduce to the corresponding probability statements on the positions of the needle of the apparatus and hence, on the position probabilities. *We, however, leave open the question whether apparatus with the postulated properties actually exist for arbitrary F and G .* For, this question essentially depends on the type of Hamiltonians that actually appear in nature. Non-relativistic quantum mechanics can make no statement regarding this point. Consequently its concepts and formalism are rather so general that the theory remains consistent for any choice of (Hermitian) Hamiltonians.

12. The W K B Approximation

(a) Limiting Transition to Classical Mechanics

The relation between the wave equation and classical mechanics has already been given in Sec. 5; namely that the mid-point of a wave-packet always moves as a point mass subject to a force, which coincides with the mean value of the classical force, taken over the wave-packet. This still does not mean a complete limiting transition to classical mechanics, since the classical force varies very sharply within the wave-packet and, therefore, the mean value of the classical force can deviate violently from its value at the position of the mid-point of the wave-packet. We obtain agreement with the properties of the system derived from the classical trajectories only if we can construct wave-packets within which the classical force varies slowly and if the wave-packet is to be considered for only such times during which the dimensions of the packet alter slightly. If we are dealing with stationary states and periodic paths, then the time referred to above should amount to many periods of revolution in order that the properties of the stationary states contained in the wave-packet, in so far as these differ relatively little among themselves, can be described approximately by means of the trajectories.

The limiting transition from quantum mechanics to classical mechanics is formally analogous to the transition from wave optics to geometrical optics (Hamilton). This analogy was indeed the starting point of the investigations of de Broglie and Schrödinger which led to the formulation of wave mechanics. This transition is made, if in the wave equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = H\psi$$

⁷ We have to construct

$$\frac{\overline{\Psi}(q, Q_n)}{(\int |\overline{\Psi}(q, Q_n)|^2 dQ)^{\frac{1}{2}}}$$

if $\overline{\Psi}(q, Q_n)$ is the function of the total system after the measurement. By the action of the apparatus all memory of the earlier states of the system is lost, since the phase of c_n will be influenced in an uncontrollable way.

we make for ψ the substitution

$$\psi = e^{\frac{i}{\hbar} S} \quad (12.1)$$

and then expand S in ascending powers of \hbar/i :¹

$$S = S_0 + \left(\frac{\hbar}{i}\right) S_1 + \left(\frac{\hbar}{i}\right)^2 S_2 + \dots \quad (12.2)$$

If we now write the Hamiltonian in cartesian co-ordinates as [cf. (5.11), (5.12); we write $V(q_1 \dots q_r)$ instead of $\sum_a V^a(x^a)$ for convenience]:

$$H = \sum_k \frac{1}{2m_k} \left(\frac{\hbar}{i} \frac{\partial}{\partial q_k} + A_k \right)^2 + V,$$

where $A_k = -(e_k/c)\Phi_k$ and V can be an arbitrary (real) function of q , then using (12.1) we have:

$$\left(\frac{\hbar}{i} \frac{\partial}{\partial q_k} + A_k \right) e^{\frac{i}{\hbar} S} = \left(\frac{\partial S}{\partial q_k} + A_k \right) e^{\frac{i}{\hbar} S},$$

$$\left(\frac{\hbar}{i} \frac{\partial}{\partial q_k} + A_k \right)^2 e^{\frac{i}{\hbar} S} = \left[\left(\frac{\partial S}{\partial q_k} + A_k \right)^2 + \frac{\hbar}{i} \left(\frac{\partial^2 S}{\partial q_k^2} + \frac{\partial A_k}{\partial q_k} \right) \right] e^{\frac{i}{\hbar} S}.$$

On separating the factor $e^{\frac{i}{\hbar} S}$ the wave equation then gives

$$-\frac{\partial S}{\partial t} = \sum_k \frac{1}{2m_k} \left[\left(\frac{\partial S}{\partial q_k} + A_k \right)^2 + \frac{\hbar}{i} \frac{\partial}{\partial q_k} \left(\frac{\partial S}{\partial q_k} + A_k \right) \right] + V. \quad (12.3)$$

On account of the series expansion (12.2), eq. (12.3) transforms finally into

$$-\frac{\partial S_0}{\partial t} = \sum_k \frac{1}{2m_k} \left(\frac{\partial S_0}{\partial q_k} + A_k \right)^2 + V = H \left(\frac{\partial S_0}{\partial q_k}, q_k \right), \quad (12.4_0)$$

i.e. in the Hamiltonian, p_k is simply replaced by $\partial S_0 / \partial q_k$. Further

$$-\frac{\partial S_1}{\partial t} = \sum_k \frac{1}{2m_k} \left[2 \left(\frac{\partial S_0}{\partial q_k} + A_k \right) \frac{\partial S_1}{\partial q_k} + \frac{\partial}{\partial q_k} \left(\frac{\partial S_0}{\partial q_k} + A_k \right) \right], \quad \left. \right\} \quad (12.4_1)$$

Instead of (12.4₁) we can also write

$$-\frac{\partial}{\partial t} e^{i\hbar S_1} = \sum_k \frac{\partial}{\partial q_k} \left[-\frac{1}{m_k} \left(\frac{\partial S_0}{\partial q_k} + A_k \right) e^{i\hbar S_1} \right]. \quad (12.5)$$

Eqs. (12.4₀) and (12.5) have a simple physical interpretation. The first equation is the well-known Hamilton-Jacobi partial differential equation of classical mechanics. In this connection it is to be remembered that the solutions of this equation in regions, where the family of classical trajectories considered lie, are real. According to (12.4₁), S_1 is also real in this region. Assuming the reality of S_0 and S_1 , (12.5) is now (neglecting S_2, \dots) identical with the *continuity equation*. In fact we have then

$$\varrho = \psi^* \psi = e^{i\hbar S_1},$$

$$i_k = \frac{1}{2m_k} \left[\psi^* \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial q_k} + A_k \psi \right) + \psi \left(-\frac{\hbar}{i} \frac{\partial \psi^*}{\partial q_k} + A_k \psi^* \right) \right] = \frac{1}{m_k} \left(\frac{\partial S_0}{\partial q_k} + A_k \right) e^{i\hbar S_1}.$$

¹ This ansatz is due to G. Wentzel, Z. Physik 38, 518 (1926), and L. Brillouin, C.R. Acad. Sci., Paris 183, 24, (1926).

Since

$$\dot{q}_k = \frac{\partial H}{\partial p_k} = \frac{1}{m_k} \left(\frac{\partial S_0}{\partial q_k} + A_k \right) \quad (12.6)$$

we have

$$i_k = \varrho \dot{q}_k, \quad (12.7)$$

and (12.5) assumes the form

$$\frac{\partial \varrho}{\partial t} + \sum_k \frac{\partial}{\partial q_k} (\varrho \dot{q}_k) = 0, \quad (12.8)$$

which corresponds to the continuity equation. If we construct, therefore, through a point in q -space at which S_0 is real, a classical trajectory according to (12.6), then the density remains constant in time along this path, according to (12.8). If in the solution S_0 of (12.4₀) no further parameter is present, it is *one* mechanical path and in the general case giving special numerical values to parameters a_1, a_2, \dots occurring in S_0 corresponds to a definite trajectory. *In the approximation considered, therefore, the wave-packets behave exactly like an aggregate of point masses, which move along classical trajectories.* The fact that these paths also satisfy the second half, viz.

$$\dot{p}_k = - \frac{\partial H}{\partial q_k}$$

of the classical equations of motion, is a simple consequence of (12.4₀) and (12.6), as is well known from the Hamilton-Jacobi theory.

The region of validity of the approximation under consideration can be characterised, according to (12.3), by requiring that the term in (12.3) multiplied by \hbar/i should be small compared to the first term; in other words, using the abbreviation

$$\pi_k = \frac{\partial S}{\partial q_k} + A_k = p_k + A_k = m_k \dot{q}, \quad (12.9)$$

we should have

$$\hbar \left| \sum_k \frac{\partial \pi_k}{\partial q_k} \right| \ll \sum_k \pi_k^2. \quad (12.10)$$

When there is no magnetic field and there is only a single particle, we have

$$\pi_k = p_k = \frac{\hbar}{\lambda} n_k,$$

where n_k represent the components of a unit vector in the direction of motion. Then (12.10) assumes the special form

$$\lambda^2 \left| \sum_k \frac{\partial}{\partial q_k} \left(\frac{n_k}{\lambda} \right) \right| \ll 1$$

or

$$\left| \sum_k \left(\lambda \frac{\partial n_k}{\partial q_k} - n_k \frac{\partial \lambda}{\partial q_k} \right) \right| \ll 1. \quad (12.10')$$

In the case of a one-dimensional problem when $n_k = \pm 1$, this inequality can be further simplified to

$$\left| \frac{\partial \lambda}{\partial x} \right| \ll 1. \quad (12.10'')$$

The inequality (12.10) is in general violated at the turning points where one of the n_k and, hence, also \dot{q}_k vanish, since $\partial n_k / \partial q_k$ can be infinitely large there. In particular this is always the case with a one-dimensional problem. In the neighbourhood of

these turning points classical mechanics, therefore, fails. In order to understand the behaviour of the solution in the neighbourhood of these critical points special investigations are required. We shall discuss this aspect presently. Equations (12.4) and (12.5) can still be used in regions which cannot be reached by classical mechanics, where S_0 will be imaginary, because in setting up these equations no assumption regarding the reality of the functions has been made.

For a stationary solution

$$\psi = e^{-\frac{i}{\hbar} Et} u$$

we have to set

$$S = -Et + \bar{S}, \quad u = e^{\frac{i}{\hbar} \bar{S}},$$

where now \bar{S} and u are independent of t . With

$$\bar{S} = \bar{S}_0 + \frac{\hbar}{i} \bar{S}_1 + \dots \quad (12.11)$$

eqs. (12.4) and (12.5) will become

$$\sum_k \frac{1}{2m_k} \left(\frac{\partial \bar{S}_0}{\partial q_k} + A_k \right)^2 + V = H \left(\frac{\partial \bar{S}_0}{\partial q_k}, q_k \right) = E, \quad (12.12_0)$$

$$0 = \sum_k \frac{1}{2m_k} \left[2 \left(\frac{\partial \bar{S}_0}{\partial q_k} + A_k \right) \frac{\partial \bar{S}_1}{\partial q_k} + \frac{\partial}{\partial q_k} \left(\frac{\partial \bar{S}_0}{\partial q_k} + A_k \right) \right], \quad (12.12_1)$$

$$0 = \sum_k \frac{\partial}{\partial q_k} \left[\frac{1}{m_k} \left(\frac{\partial \bar{S}_0}{\partial q_k} + A_k \right) e^{i\bar{S}_1} \right]. \quad (12.13)$$

It is easy to generalise the above considerations to the case of curvilinear co-ordinates as follows. If

$$ds^2 = \sum_x \sum_{\lambda} g_{x\lambda} dq_x dq_{\lambda}$$

is the line element, $g^{x\lambda}$ the matrix reciprocal to $g_{x\lambda}$ ($g_{x\lambda} g^{x\lambda} = \delta_x^{\lambda}$), and D is equal to the square root of the determinant $g = |g_{x\lambda}|$, then according to eq. (5.40) the wave equation reads:

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \frac{1}{2D} \left(\frac{\hbar}{i} \frac{\partial}{\partial q_x} + A_x \right) D g^{x\lambda} \left(\frac{\hbar}{i} \frac{\partial}{\partial q_{\lambda}} + A_{\lambda} \right) \psi + V \psi.$$

Here the factors m_x have been absorbed in $g_{x\lambda}$. A summation is understood wherever an index appears twice. The substitution of (12.1) in the above equation then gives,

$$-\frac{\partial S}{\partial t} = \frac{1}{2} g^{x\lambda} \left(\frac{\partial S}{\partial q_x} + A_x \right) \left(\frac{\partial S}{\partial q_{\lambda}} + A_{\lambda} \right) + \frac{1}{2} \frac{1}{D} \frac{\hbar}{i} \frac{\partial}{\partial q_x} D g^{x\lambda} \left(\frac{\partial S}{\partial q_{\lambda}} + A_{\lambda} \right) + V \quad (12.3')$$

in the place of (12.3). Insertion of the expansion (12.2) with (12.3') leads (remembering that $g^{x\lambda} = g^{x\lambda}$) to:

$$-\frac{\partial S_0}{\partial t} = \frac{1}{2} g^{x\lambda} \left(\frac{\partial S_0}{\partial q_x} + A_x \right) \left(\frac{\partial S_0}{\partial q_{\lambda}} + A_{\lambda} \right) + V = H \left(\frac{\partial S_0}{\partial q_x}, q_x \right), \quad (12.4_0)$$

$$-\frac{\partial S_1}{\partial t} = g^{x\lambda} \left(\frac{\partial S_0}{\partial q_x} + A_x \right) \frac{\partial S_1}{\partial q_{\lambda}} + \frac{1}{2} \frac{1}{D} \frac{\hbar}{i} \frac{\partial}{\partial q_x} D g^{x\lambda} \left(\frac{\partial S_0}{\partial q_{\lambda}} + A_{\lambda} \right) \quad (12.4_1)$$

or

$$-\frac{\partial}{\partial t} e^{i\bar{S}_1} = \frac{1}{D} \frac{\partial}{\partial q_x} \left[D g^{x\lambda} \left(\frac{\partial S_0}{\partial q_{\lambda}} + A_{\lambda} \right) e^{i\bar{S}_1} \right]. \quad (12.5')$$

For the probability and current density,

$$p = D \psi \psi^*, \quad i^* = D g^{x\lambda} \left[\psi^* \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial q_x} + A_x \psi \right) + \psi \left(-\frac{\hbar}{i} \frac{\partial \psi^*}{\partial q_x} + A_x \psi^* \right) \right]$$

we have, in general, from the wave equation the continuity equation

$$\frac{\partial \varrho}{\partial t} + \frac{\partial}{\partial q_x} i^* = 0.$$

For real S_0 and S_1 , again in the approximation considered, we have

$$\varrho = e^{2S_1},$$

$$i^* = \frac{\partial H}{\partial p_x} = g^{*1} \left(\frac{\partial S_0}{\partial q_1} + A_1 \right). \quad (12.6')$$

$$i^* = \varrho \dot{q}^*.$$

$$(12.7')$$

From these equations the same results can be derived as was done in the cartesian co-ordinates earlier.

The introduction of curvilinear co-ordinates is important because by suitable choice of these co-ordinates for special Hamiltonians, a separation of variables can often be achieved. In this case, the stationary solution u can be factorised into functions of individual variables

$$u = u_1(q_1) \dots u_f(q_f)$$

and correspondingly \bar{S} decomposes additively as follows:

$$\bar{S} = \bar{S}_1(q_1) + \dots + \bar{S}_f(q_f).$$

Each of these functions $\bar{S}_i(q_i)$ then satisfies a differential equation of the second order in this variable q_i in which, however, $f-1$ constants a_1, \dots, a_f appear as parameters in addition to the (constant) energy. All that is said in the following regarding systems of a single degree of freedom, holds, mutatis mutandis, also for the motion represented by each of the separated co-ordinates q_i and the associated eigenfunction $u_i(q_i)$ of a separable system. In particular, this is the case for the radial motion of a point mass under the influence of a central force.

We shall now consider the *one-dimensional* case more precisely. Here it is essential to set the vector potential equal to zero since it does not give rise to a magnetic field in the one-dimensional case. If we desire to have the stationary solution and if we now write for the sake of simplicity S instead of \bar{S}_0 and 'a' instead of $e^{\bar{S}}$, so that in the approximation considered

$$u = a e^{\frac{i}{\hbar} S},$$

then (12.12₀) and (12.13) will become

$$\frac{1}{2m} \left(\frac{dS}{dx} \right)^2 + V(x) = E \quad (12.14)$$

and

$$0 = \frac{d}{dx} \left(a^2 \frac{dS}{dx} \right). \quad (12.15)$$

Therefore, with

$$p(x) = \sqrt{2m(E - V(x))} = \pm \frac{ds}{dx} \quad (12.16)$$

we get

$$S = \pm \int p(x) dx.$$

Here we can suitably choose the lower limit of integration. The two signs correspond to two different particular solutions of eq. (12.14). It then follows that

$$a^2 \frac{dS}{dx} = \text{const} = c^2,$$

$$a = \frac{C}{\sqrt{p(x)}} = \frac{C}{\sqrt{2m(E - V(x))}}. \quad (12.17)$$

The general solution of (12.14) and (12.15), therefore, reads

$$u = \frac{C_1}{\sqrt{|p(x)|}} e^{-\frac{i}{\hbar} \int_a^x |p(s)| ds} + \frac{C_2}{\sqrt{|p(x)|}} e^{-\frac{i}{\hbar} \int_x^a |p(s)| ds}. \quad (12.18)$$

This solution fails in the neighbourhood of the points where $p(x)$ vanishes. Even in regions where $E - V(x)$ is negative (and $p(x)$ is pure imaginary) and hence can never be reached by the classical trajectory, a solution of the form (12.18) exists. For the sake of clarity, we write it for such a region in the form

$$u = \frac{C'_1}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar} \int_a^x |p(s)| ds} + \frac{C'_2}{\sqrt{|p(x)|}} e^{+\frac{1}{\hbar} \int_x^a |p(s)| ds}. \quad (12.18')$$

The consideration of differential equations (12.14) and (12.15) of geometrical optics (classical mechanics) alone is not sufficient to connect correctly the particular solutions (12.18) and (12.18') at the critical points which coincide with the turning points for $p(x) = 0$. Since here we have to understand by the word "correct" that these different particular solutions of eqs. (12.14) and (12.15) should approximate to the same particular solution $u(x)$ of the wave equation

$$\hbar^2 \frac{d^2 u}{dx^2} + p^2(x) u = 0$$

in the regions where $p^2(x) > 0$ and $p^2(x) < 0$, it is absolutely necessary to study the wave equation at least in the neighbourhood of the turning points.

The problem raised above, which is not mathematically quite simple, has been investigated in detail by Kramers and his students.² The result is as follows. We assume that $V(x)$ is continuous at a turning point. (This enables us to use the complex plane in the proof.) Let to the right of the turning point at $x = a$, i.e. for $x > a$, $p^2(x) > 0$ and for $x < a$, $p^2(x) < 0$. Then according to Kramers,

$$\left. \begin{aligned} u(x) &= \frac{C}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar} \int_a^x |p(s)| ds} && \text{for } x < a, \\ u(x) &= \frac{C}{\sqrt{|p(x)|}} 2 \cos \left[\frac{1}{\hbar} \int_a^x |p(s)| ds - \frac{\pi}{4} \right] && \text{for } x > a \end{aligned} \right\} \quad (12.19)$$

is a "correct" connection. More generally

$$\left. \begin{aligned} u(x) &= \frac{C}{i\sqrt{|p(x)|}} e^{+\frac{1}{\hbar} \int_a^x |p(s)| ds} + \frac{1}{2} \frac{C}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar} \int_a^x |p(s)| ds} && \text{for } x < a, \\ u(x) &= \frac{C}{\sqrt{|p(x)|}} e^{i \left[\frac{1}{\hbar} \int_a^x |p(s)| ds - \frac{\pi}{4} \right]} && \text{for } x > a. \end{aligned} \right\} \quad (12.20)$$

² H.A. Kramers, Z. Physik **39**, 828 (1926). – A. Zwaan, Dissert. Utrecht (1929), see in particular Chapter III, Sec. 2. – K.F. Niessen, Ann. d. Phys. **85**, 497 (1928). – H.A. Kramers and G.P. Ittmann, Z. Physik **58**, 217 (1929), in particular pp. 221 and 222.*

* See also R.E. Langer, Phys. Rev. **51**, 669 (1937).

If we construct the real part of both the expressions in (12.20), which is permitted, then we again get back the connection (12.19). Also by taking the imaginary parts we get another correct association. It is no doubt true that the second term in the first expression (12.20) is so small when compared with the first that it lies within the limits of error in the asymptotic representation and can be neglected for practical purposes. In all problems in which not only the magnitude of $u(x)$ but also its phase plays a role, we are justified in retaining the second term. In particular this retention will take care that the current, for $x < a$, will be equal to the current, for $x > a$, as is required by the continuity equation.

At a second turning point $b > a$, which has the allowed region to the left, we have a connection similar to (12.19):

$$\left. \begin{aligned} u(x) &= \frac{C'}{\sqrt{|\rho(x)|}} e^{-\frac{1}{\hbar} \int_a^x |\rho(x)| dx} && \text{for } x > b, \\ u(x) &= \frac{C'}{\sqrt{\rho(x)}} 2 \cos \left[\frac{1}{\hbar} \int_x^b \rho(x) dx - \frac{\pi}{4} \right] && \text{for } x < b. \end{aligned} \right\} \quad (12.20')$$

We can now give the quantum conditions if the region in which $p^2(x) > 0$ consists of a single, continuous interval $a < x < b$ for the concerned energy value E . We introduce this assumption explicitly. Then in order that $u(x)$ remains bounded for $x = -\infty$ as well as for $x = +\infty$, the particular solutions which increase exponentially in both the regions $x < a$ and $x > b$ must drop out and the particular solutions steeply falling there alone survive. For this purpose, it is necessary, according to eqs. (12.20) and (12.20') that for all x in the interval $a < x < b$

$$C \cos \left[\frac{1}{\hbar} \int_a^x \rho(x) dx - \frac{\pi}{4} \right] = C' \cos \left[\frac{1}{\hbar} \int_x^b \rho(x) dx - \frac{\pi}{4} \right].$$

This is possible only if the (constant) sum of the two phases is equal to a multiple of π :

$$\frac{1}{\hbar} \int_a^b \rho(x) dx - \frac{\pi}{2} = n\pi,$$

or with

$$J = 2 \int_a^b \rho(x) dx = \phi \rho(x) dx, \quad (12.21)$$

if

$$\frac{1}{\hbar} J = \left(n + \frac{1}{2} \right) 2\pi, \quad J = \left(n + \frac{1}{2} \right) \hbar. \quad (12.22)$$

Further

$$C' = (-1)^n C.$$

This agrees with the statement that n signifies the number of nodes of the eigenfunction. In fact the phase

$$\frac{1}{\hbar} \int_a^x \rho(x) dx - \frac{\pi}{4}$$

increases from $-\frac{1}{4}\pi$ to $(n + \frac{1}{2})\pi - \frac{1}{4}\pi$, if x goes from a to b ; therefore, the cosine assumes the value zero n -times, in this interval.

(b) Relation to the Old Quantum Theory

The result (12.22) leads to the quantisation rule of the old quantum theory but with half-integral quantum numbers, i.e. the phase integral of the old quantum theory will be a half-integral multiple of \hbar . This represents a better approximation of quantum mechanics than the quantisation with integral quantum numbers. As mentioned earlier, this holds also for systems with many degrees of freedom if the variables are separable. In this, it is, however, assumed that the degree of freedom concerned is of *oscillatory type*, i.e. that to each value of the co-ordinate q concerned belong, in a particular interval, two values of the velocity \dot{q} of the particle, which differ in sign, so that each point will be passed through twice in the course of a complete period. The q -point outside the interval concerned can never be reached by any trajectory with the same values of the integration constant. The oscillatory type of motion with a single degree of freedom stands in contrast to the *rotational type* which, e.g., is present in the case of cyclical angle variables (precessional motion about a spatially fixed axis). We shall see that in this case (so long as we are dealing with the rotational motion of a particle and not its spin) quantum mechanics leads to integral quantisation.

If more than one interval is present, which can be reached by classical paths with the same total energy, then according to quantum mechanics, characteristic effects occur, which depend on the fact that in the classically forbidden regions the wave function is not exactly zero but only very small, as follows from eq. (12.20). While classically two regions can be completely separated out, for a particle of fixed total energy, by a "potential barrier" of finite height, in quantum mechanics the wave function can "tunnel through" from one region to the other and the stationary solution will have similar densities in both the regions.³ This fact is of fundamental significance for many applications of quantum theory.⁴ We shall meet with this situation again in Sec. 24.

The condition (12.10") is satisfied the better at the turning point, the larger the quantum number n is. Remembering that for large quantum numbers, the phase

$$\frac{1}{\hbar} \int_a^x p(x) dx - \frac{\pi}{4}$$

is a rapidly varying function, we can now normalise our eigenfunctions. Such integrals containing a rapidly changing phase are to be neglected compared to the others which contain only a slowly oscillating or non-oscillating phase. In this approximation, the eigenfunctions are also orthogonal. For normalisation, we obtain the condition

$$4C^2 \int_a^b \frac{dx}{p(x)} \cos^2 \left[\frac{1}{\hbar} \int_a^x p(x) dx - \frac{\pi}{4} \right] = 1$$

or

$$2C^2 \int_a^b \frac{dx}{p(x)} = 2C^2 \int_a^b \frac{dx}{m\dot{x}} = \frac{C^2}{m\omega} = 1,$$

³ A direct demonstration of the presence of a particle on the potential hill by measurement of position is always connected with such an indeterminacy in the energy supplied to the particle that after the supply of this energy the particle can overcome the potential barrier even classically.

⁴ Cf. for this especially E. Schrödinger, Berl. Ber. (1929) 668.

if [remembering the definition (12.21) of the phase integral J]

$$\frac{1}{\omega} = 2 \int_a^b \frac{dx}{\dot{x}} = 2m \int_a^b \frac{dx}{\sqrt{2m(E - V(x))}} = \frac{\partial J}{\partial E} \quad (12.23)$$

denotes the period of the classical motion. Therefore

$$u_n(x) = \sqrt{\frac{m\omega_n}{p_n(x)}} 2 \cos \left[\frac{1}{\hbar} \int_{a_n}^x p_n(x) dx - \frac{\pi}{4} \right] \quad (12.24)$$

will be the normalised eigenfunction. We now construct the matrix x_{nm}

$$x_{nm} = \int x u_n u_m dx = \int x \sqrt{\frac{m\omega_n}{p_n(x)}} \sqrt{\frac{m\omega_m}{p_m(x)}} 2 \left\{ \cos \left[\frac{1}{\hbar} \int_{a_n}^x p_n(x) dx - \frac{1}{\hbar} \int_{a_m}^x p_m(x) dx \right] + \cos \left[\frac{1}{\hbar} \int_{a_n}^x p_n(x) dx + \frac{1}{\hbar} \int_{a_m}^x p_m(x) dx - \frac{\pi}{2} \right] \right\} dx.$$

Since the second term here containing the sum of the phases oscillates more rapidly than the first term, which contains the difference of the phases, the second term can be neglected compared to the first. Further, all quantities vary relatively slowly and smoothly with the quantum number, so that we can replace all the differences of the form $F_n - F_m$ by differential coefficients $(n - m)\hbar \frac{\partial F}{\partial J}$, if J again denotes the phase integral defined by eq. (12.21). The differentiation with respect to the lower limit of the integral $\int_a^x p(x) dx$ gives no contribution, since the integrand vanishes there. As Debye⁵ has shown, on setting

$$\tau = n - m \quad (12.25)$$

we get in this way,

$$x_{nm} = 2 \int_a^b x \frac{m\omega}{p(x)} \cos \left(2\pi\tau \int_a^x \frac{\partial p}{\partial J} dx \right) dx.$$

Now

$$\frac{\partial p}{\partial J} = \frac{\partial p}{\partial E} / \frac{\partial J}{\partial E} = \omega \frac{m}{p(x)} = \omega/\dot{x}.$$

Therefore

$$\int_a^x \frac{\partial p}{\partial J} dx = \omega t$$

and

$$x_{nm} = 2 \int_0^{1/\omega} x(t) \omega \cos(2\pi\tau\omega t) dt = \omega \int_0^{1/\omega} x(t) \cos(2\pi\tau\omega t) dt.$$

⁵ P. Debye, Phys. Z. 28, 170 (1927).

If we decompose the classical motion into Fourier components, in which at time $t = 0$, $x = a$, then

$$x = \sum_{\tau=0}^{\infty} a_{\tau} \cos 2\pi \tau \omega t$$

and

$$a_{\tau} = 2\omega \int_0^{1/\omega} x(t) \cos(2\pi\omega\tau t) dt.$$

We, therefore, obtain the relation

$$x_{nm} = \frac{1}{2} a_{\tau} = \frac{1}{2} a_{n-m} \quad (12.26)$$

in the limit of large quantum numbers. The factor $\frac{1}{2}$ is correct, since the sum

$$x_{n,n+\tau} e^{2\pi i v_{n,n+\tau} t} + x_{n,n-\tau} e^{2\pi i v_{n,n-\tau} t}$$

is exactly equal to

$$a_{\tau} \cos 2\pi \tau \omega t.$$

Similarly we can calculate the momentum matrix instead of the co-ordinate matrix. Through these results the connection between the original form of Bohr's correspondence principle and the above theory is established.

By collecting a number of eigenfunctions to form a group we can easily build up a wave-packet in the region of large quantum numbers, which traverses a period in the neighbourhood of the classical trajectory. It represents a state in which a particle trajectory exists in an approximate sense.^{6,7} If the packet covers the states from $n - k$ to $n + k$, then the time t at which the packet passes through a particular place is inaccurate by Δt , given by

$$\omega \Delta t \sim \frac{1}{h}$$

Since $\Delta E = \frac{\partial E}{\partial J} \Delta J = \omega k h$, we have

$$\Delta E \Delta t \sim h \quad (12.27)$$

On the other hand, if $\omega t + \delta_0 = w$ is introduced as the angle variable, then

$$\Delta J \Delta w \sim h. \quad (12.27')$$

Here it is assumed that the wave-packet covers a large number of states since otherwise the phase w completely loses its sense. It should be mentioned that it is possible in systems of the kind considered here to define not w as such, but e^{iw} as an operator (function of p and q) and similarly⁸ for J . Here J is Hermitian and e^{iw} unitary. The two operators satisfy the C.R.

$$J e^{iw} - e^{iw} J = e^{iw} h \quad \text{or} \quad J e^{iw} = e^{iw} (J + h), \quad (12.28a)$$

from which, by left-multiplication with e^{-iw} and again right-multiplication with e^{-iw} it follows that

$$e^{-iw} J - J e^{-iw} = e^{-iw} h \quad \text{or} \quad J e^{-iw} = e^{-iw} (J - h). \quad (12.28b)$$

⁶ P. Debye, I.c. footnote 5.

⁷ C.G. Darwin, Proc. Roy. Soc. Lond., Ser. A 117, 258 (1927), in particular § 8.

⁸ P.A.M. Dirac, Proc. Roy. Soc. Lond., Ser. A 111, 279 (1926).

However, we shall not need these operators for applications.

It is essential that the knowledge of the phase of revolution of the particle and that of the stationary state be mutually exclusive. The phase of revolution (path) of a particle in a single stationary state does not exist, since every experiment to determine it throws the system into another stationary state. A temporal development of the density of a group of waves formed from many stationary states is similar to a classical path only in the limiting case of large quantum numbers. This follows from the fact that a wave-packet occupies, as a consequence of the uncertainty relation, at least the area \hbar in phase space. On the other hand, the classical trajectory with energy E_n corresponding to the n^{th} state encloses a definite area $n\hbar$ in the phase space. Only if n is large will this area be large compared to \hbar so that only then will an effectively orbiting density packet be possible.

We have still to discuss as to how a wave-packet of the kind considered above behaves in course of a long time. For this it is essential that the frequencies

$$\frac{E_{n+\tau} - E_n}{\hbar},$$

(where $-k \leq \tau \leq +k$, $-k \leq \sigma \leq +k$) which occur in the density of the packet are not, in general, exact multiples, $(\tau - \sigma)\omega$, of a fundamental frequency, as is the case in the classical periodic trajectory. This is due to the fact that the classical frequency ω depends, in general, on the value of the phase integral. We can set approximately

$$\frac{1}{\hbar} (E_{n+\tau} - E_n) = \tau \omega + \frac{1}{2} \hbar \frac{\partial \omega}{\partial J} \tau^2,$$

$$\frac{1}{\hbar} (E_{n+\tau} - E_{n+\sigma}) = (\tau - \sigma) \omega + \frac{1}{2} \hbar \frac{\partial \omega}{\partial J} (\tau^2 - \sigma^2).$$

As Darwin has shown, this has the consequence that the indeterminacy $1/k$ in the phase is stepped up by a further indeterminacy $\hbar k \frac{\partial \omega}{\partial J} t$. After time

$$\Delta t \sim \frac{1}{\hbar \hbar \frac{\partial \omega}{\partial J}} \quad (12.29)$$

the packet is, therefore, smeared over the whole circumference of the trajectory and the phase is completely lost. The number N of the periods after which this happens, is given by

$$N = \omega \Delta t \sim \frac{\omega}{\hbar \hbar \frac{\partial \omega}{\partial J}}. \quad (12.29')$$

Only in special systems, as e.g., the harmonic oscillator or the motion of an electron in a plane perpendicular to a homogeneous magnetic field, whose frequency ω is strictly independent of the initial conditions, does the wave-packet stay on together permanently. In the first case, Schrödinger⁹ and in the second case Kennard¹⁰ and Darwin¹¹ have given rigorous solutions for wave groups of this nature.

Finally we may refer to the fact that the old quantum theory succeeded in making

⁹ E. Schrödinger, Naturwiss 14, 664 (1926).

¹⁰ E.H. Kennard, Z. Physik 44, 326 (1927).

¹¹ C.G. Darwin, I.c. footnote 7.

statements on stationary states only for the special case of multiple-periodic systems whereas the eigenvalue problem of quantum mechanics always possesses a solution. Even in the general case, it is always possible, as we have seen, to construct wave-packets which move along the classical trajectories. Due to the passage of time, however, a spreading of the wave-packet always sets in. Therefore, the fact that only discrete frequencies are present in the density of such a packet is not connected directly with the simple periodic behaviour of the classical trajectories over long periods of time. Now the periodic behaviour of the classical paths for long periods of time in, e.g., the three body problem is quite complicated but this fact is not of any significance for applications to atomic systems, and this is a progress made by wave mechanics.

CHAPTER VII

Identical Particles, Spin and Exclusion Principle

13. Hamiltonian Functions with Transformation Groups. Angular Momentum and Spin¹

(a) Group Theoretical Considerations

If the Hamiltonian operator is invariant under a certain group of transformations of variables, it follows that from a solution $u_n(q)$ of the wave equation, new solutions (of the wave equation) can be obtained by performing the transformations of the group on the original solution. If T is a transformation of the group, H , the Hamiltonian, f an arbitrary function and $u(q)$ a solution of the equation

$$H u(q) = E u(q),$$

it follows from the validity of the relation

$$T(Hf) = H(Tf)$$

for all f , that the function

$$v(q) = T u(q)$$

satisfies the equation

$$H v(q) = E v(q).$$

If to the energy eigenvalue E belong a finite number, say h , of eigenfunctions (h -fold degeneracy) then in the h -dimensional vector space which corresponds to the eigenvalue E , there is a basis u_1, u_2, \dots, u_h , such that each solution $v(q)$ of the equation

$$H v = E v$$

¹ On the relation between quantum mechanics and group theory there exist several detailed textbooks: H. Weyl, The Theory of Groups and Quantum Mechanics, Dover Publications, New York (1950); E.P. Wigner, Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra, Academic Press, New York (1959); B.L. van der Waerden, Die Gruppentheoretische Methode in der Quantenmechanik, Springer Verlag, Berlin (1932); W. Pauli, Continuous Groups in Quantum Mechanics, Mimeographed lectures, CERN-publications, (1956).* We make here a very rapid survey of the subject and refer the reader for all proofs and details to these textbooks.

* Now included as an article in Springer Tracts in Modern Physics, 37 (1965); See also the article, Group Theory and Spectroscopy by G. Racah in the same volume; M. Hamermesh: Group Theory and Its Applications to Physical Problems, Addison-Wesley, Reading, Mass., U.S.A. (1962).

can be represented in the form

$$v = \sum_k c_k u_k.$$

The transformations T of the group, when applied to u_1, \dots, u_k , transform, therefore, this vector space *linearly*, such that to the sequence of two different transformations T corresponds the sequence of the associated linear mappings. In order to be in accord with the law of matrix multiplication, it is convenient to stipulate the following. If we carry out on the variables q of the function $u(q)$ the transformation T , defined by $q' = f(q_1 \dots q_r)$, then we associate this transformation of the variables q with an operator T , which changes the function $u(q_1 \dots q_r)$ to $u'(q_1 \dots q_r)$:

$$T u \equiv u'.$$

Here $u'(q_1 \dots q_r) = u(q_1 \dots q_r)$, so that

$$u'(T q) = u(q)$$

or

$$u'(q) = u(T^{-1} q)$$

Only in this case is the composition of two operators in the order, first T_2 and then T_1 , next, associated with *the same* sequence of transformation of the variables q . In fact

$$T_2 u(q) = u'(q) = u(T_1^{-1} q)$$

and if we replace q by $T_1^{-1} q$, then we obtain

$$(T_1 T_2) u = u'(T_1^{-1} q) = u(T_1^{-1} T_2^{-1} q) = u((T_1 T_2)^{-1} q).$$

In matrix form we have to set

$$(T u_i) = \sum_{k=1}^K u_k c_{ki}. \quad (13.1)$$

Then we have

$$c_{ki}(T_1 T_2) = \sum_m c_{km}(T_1) c_{mi}(T_2). \quad (13.2)$$

or in matrix form

$$c(T_1 T_2) = c(T_1) c(T_2). \quad (13.2')$$

We then say that the associated linear mapping forms a *representation* of the group. Naturally *different* transformations of the group can correspond to the same linear mapping in the representation. If the Jacobian of the new variables $q'_r = f_r(q_1 \dots q_r)$ which are defined by the transformation T has the value 1 with respect to the old q_r and if further the q'_r as well as the q_r are real, it then follows, with

$$v(q_1 \dots q_r) = T v(f_1(q), f_2(q) \dots f_r(q)),$$

that

$$\int v_h^* v_i dq = \int (T v_h)^* (T v_i) dq.$$

In this case the matrices $c(T)$ are unitary for all T ; we also call the representation unitary. In this case a normalised orthogonal system will be transformed to a normalised orthogonal system again.

The concept of the *reduction* of a representation is very important. A representation is said to be *reducible*, if an invariant sub-space of smaller dimension

than the original representation-space exists. That is, for a suitable choice of the basis, the linearly independent functions u_1, \dots, u_g ($g < h$) which form only a part of the complete basis $u_1 \dots u_h$, transform into themselves under the transformation T . The total matrix $c(T)$ has for this choice of the basis, the form

$$c = \begin{pmatrix} a & r \\ 0 & b \end{pmatrix} \quad (13.3)$$

where a is g -dimensional and b is $(h - g)$ -dimensional. If there is no proper invariant subspace, the representation is said to be *irreducible*. A change of the basis means a transformation $c' = ScS^{-1}$ of the representation matrix. Two representations which are related in this way, are called equivalent. If c is a unitary matrix, then from the form (13.3) of the matrix, it readily follows that by changing the basis, even r can be made to vanish, i.e. the representation c is *decomposable*. For a finite group it can be proved that every representation is equivalent to a unitary representation and that, therefore, every reducible representation is decomposable. This is not always the case for continuous groups, but is true only for a particular class of these groups, the semi-simple groups. The rotation group is one such group as well as the group of all linear transformations with determinant 1 (where the unimodularity is important). Since we have to do only with unitary representations in quantum mechanics, we need not go into the complications of the general case here.

Thus every representation (D) of a group decomposes into irreducible representations according to

$$(D) = (D_1) + (D_2) + \dots,$$

and indeed it can be shown that this decomposition is unique. The degeneracy that corresponds to the order of the irreducible representation, which belongs to the energy eigenvalue E , cannot be removed by a continuous change of the Hamiltonian so long as the latter is invariant under the group concerned (in contrast to the accidental degeneracy which corresponds to the higher order of a reducible representation). If we, however, alter the Hamiltonian such that it is only invariant under a sub-group of the original group, then under this smaller group the representation will in general be reducible. The change of the basis chosen in order to decompose the representation, which we can also call the reduction of the original representation with respect to the sub-group, corresponds, in general, to the splitting of the original energy value E into different eigenvalues, when a perturbation, which is invariant only under this sub-group is applied.

From two representations (D_1) and (D_2) of degree h_1 and h_2 respectively, we can construct the product representation $(D_1 \times D_2)$ of degree $h_1 \cdot h_2$ in the following way. From the basis u_k ($k = 1, 2, \dots, h_1$) of (D_1) and v_l ($l = 1, 2, \dots, h_2$) of (D_2) we build the $h_1 \cdot h_2$ products $u_k v_l$. If the u_k undergoes a linear transformation $D_1(T)$ and v_l a linear transformation $D_2(T)$ then $u_k v_l$ also undergoes a linear transformation and this will be defined as $(D_1 \times D_2)$. In particular we can take $(D_1) = (D_2)$. Naturally $(D_1 \times D_2)$ is reducible in general even if (D_1) and (D_2) are irreducible. By changing the basis of the $h_1 h_2$ -dimensional space we can, therefore, reduce $(D_1 \times D_2)$ in which the irreducible parts of (D_1) and (D_2) can be different. This direct product form of representation always occurs in the couplings of independent systems.

For the continuous groups, the infinitesimal transformations which lie in the neighbourhood of the identity element are of special interest. For, these by

themselves form a linear manifold – a vector space of as many dimensions as the group has independent parameters. (For the group of rotations in three-dimensional space the linear manifold is, therefore, a three-dimensional vector space.) In fact it follows from $T(0, \dots, 0) = 1$ that

$$T(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_r) = 1 + \varepsilon_1 \omega_1 + \varepsilon_2 \omega_2 + \dots + \varepsilon_r \omega_r,$$

if T depends on ε continuously. The $\omega_1, \omega_2, \dots, \omega_r$ are, therefore, similar to the T operators which act on the variables q of the eigenfunctions. The fact, that the transformations concerned generate a group, requires that the “commutator” $[\omega_p, \omega_q] = \omega_p \omega_q - \omega_q \omega_p$ can be expressed in terms of the ω themselves, with coefficients which are characteristic of the group concerned:

$$[\omega_p, \omega_q] = \sum_{s=1}^r c_{pq,s} \omega_s. \quad (13.4)$$

These coefficients need satisfy only certain relations, which arise from the identity*

$$[[\omega_p, \omega_q] \omega_r] + [[\omega_q, \omega_r] \omega_p] + [[\omega_r, \omega_p] \omega_q] \equiv 0.$$

In the stipulation made above regarding the association of the operators acting on the eigenfunctions with the transformations of the variables, both (the transformations and the corresponding operators) obey the same C.R., and the two need not be distinguished from each other. The fact that the Hamiltonian is invariant under the group concerned is expressed by saying that ω and H commute:

$$\omega_p H - H \omega_p = 0.$$

This is the case also for the operator T of finite transformation. Equivalently we can say that the ω_p , as matrices, are constant in time, i.e. they are integrals of the equations of motion. Up to a factor i , the ω_p are Hermitian, if the operators T are unitary (cf. Sec. 10). To every representation of the group belongs, in general, a system of matrices for the ω_k which satisfy the relation (13.4).

As an illustration of the above statements, let us consider the group of translations which displaces the co-ordinates $x_k^{(a)}$, ($k = 1, 2, 3$) by A_k

$$x_k'^{(a)} = x_k^{(a)} + A_k, \quad (A_1, A_2, A_3 \text{ continuous parameters}); \quad \omega_k = \sum_a \frac{\partial}{\partial x_k^{(a)}}.$$

This corresponds, up to a factor \hbar/i , to the total momentum of the system. In fact, the invariance of the Hamiltonian under this group of transformations can be equivalently stated as follows: the potential energy depends only on the relative co-ordinates of the particle.

We now proceed to consider the group of rotations of the space co-ordinates simultaneously for all particles of the system. Here we can employ two different methods. Either we start from the infinitesimal standpoint, find out the form of the operators ω_k belonging to the *infinitesimal* rotations and, on the basis of their C.R., obtain the associated representation-matrices in a purely algebraic way; or, we try, using the method of analysis, to find the representations belonging to the finite rotations. Each method supplements the other. Let us begin with the first method.

*Called the Jacobi identity. The ω 's satisfying this identity and eq. (13.4) are said to form a Lie Algebra. The corresponding Lie group has T as the elements in the neighbourhood of the identity. The coefficients $c_{pq,s}$ are called structure constants.

As the three independent infinitesimal rotations of the three-dimensional space, we choose the rotations about the co-ordinate axes

$$\delta x_1 = 0, \quad \delta x_2 = -\varepsilon_1 x_3, \quad \delta x_3 = +\varepsilon_1 x_2, \quad (13.5)$$

where the two remaining infinitesimal rotations are obtained by cyclical permutation. On the basis of an elementary kinematic consideration (limiting transition from finite to infinitesimal rotations), we obtain the C.R. characteristic of infinitesimal rotations:

$$\omega_1 \omega_2 - \omega_2 \omega_1 = \omega_3, \dots, \quad (13.6)$$

Here the operators ω or the corresponding linear mappings are so defined that, for example, the operator $1 + \varepsilon_1 \omega_1$ corresponds to the transformation (13.5). These relations must then be satisfied by all representations of the rotation group. Since we shall investigate also reflections of the space co-ordinates later, it is to be specially mentioned that the ω behave like an antisymmetric tensor and not like a vector under rotation. If we, therefore, write $\omega_{23}, \omega_{31}, \omega_{12}$ with $\omega_{ik} = -\omega_{ki}$ instead of $\omega_1, \omega_2, \omega_3$, then (13.6) can be written as

$$\omega_{ik} \omega_{lm} - \omega_{im} \omega_{lk} = -\delta_{kl} \omega_{im} - \delta_{im} \omega_{lk} + \delta_{il} \omega_{km} + \delta_{km} \omega_{il} \quad (13.6')$$

($\delta_{ik} = 0$ for $i \neq k$ and $=1$ for $i = k$). This form of C.R. for the infinitesimal rotations is also valid in an n -dimensional space. We shall make use of this property in the study of the Lorentz group later.

According to our earlier requirement, to the infinitesimal rotation (13.5) belongs, for a single particle, the operator $1 + \varepsilon \omega_1$ (or $1 + \varepsilon \omega_2$) which takes $u(x_1 x_2 x_3)$ to $u(x_1, x_2 + \varepsilon x_3, x_3 - \varepsilon x_2)$ and, therefore,

$$\omega_1 u = -\left(x_2 \frac{\partial u}{\partial x_3} - x_3 \frac{\partial u}{\partial x_2} \right).$$

In the many-particle case the co-ordinates of all the particles must be subject to the same rotation and so we obtain

$$\omega_k u = -\sum_r \left(x_2^{(r)} \frac{\partial u}{\partial x_3^{(r)}} - x_3^{(r)} \frac{\partial u}{\partial x_2^{(r)}} \right),$$

where we have to sum over all the particles involved. Since the linear momentum $p_k^{(r)}$ will be represented by the operator $\frac{\hbar}{i} \frac{\partial}{\partial x_k^{(r)}}$, the ω are connected to the total angular momentum

$$\mathbf{P}_i = \sum_r \mathbf{x}_2^{(r)} \mathbf{p}_3^{(r)} - \mathbf{x}_3^{(r)} \mathbf{p}_2^{(r)} = \frac{\hbar}{i} \sum_r \left(x_2^{(r)} \frac{\partial}{\partial x_3^{(r)}} - x_3^{(r)} \frac{\partial}{\partial x_2^{(r)}} \right), \quad (13.7)$$

by the simple relation

$$\omega_k = -\frac{i}{\hbar} \mathbf{P}_k. \quad (13.8)$$

Again we can verify that ω is i times a Hermitian operator. In fact it follows from the C.R. (5.21) for p_k and q_k that

$$\mathbf{P}_1 \mathbf{P}_2 - \mathbf{P}_2 \mathbf{P}_1 = -\frac{\hbar}{i} \mathbf{P}_3. \quad (13.9)$$

It is, however, of importance here to show that the existence of the integrals of the three components of the angular momentum which, up to a pure imaginary factor, coincide with the operators ω belonging to infinitesimal rotation follows

independently from the C.R. (5.21). The C.R. of ω themselves follow directly from the kinematics of the rotation group. This statement holds also for the C.R.

$$[\omega_k, \mathbf{C}] = 0, \quad [\mathbf{P}_k, \mathbf{C}] = 0 \quad (13.10)$$

for each scalar operator \mathbf{C} and

$$[\omega_1, \mathbf{A}_3] = -[\omega_3, \mathbf{A}_1] = \mathbf{A}_3, \quad (13.11)$$

$$[\mathbf{P}_1, \mathbf{A}_3] = -[\mathbf{P}_3, \mathbf{A}_1] = -\frac{\hbar}{i} \mathbf{A}_3 \quad (13.11')$$

for the components of every vector operator \vec{A} . Here it is assumed that \mathbf{C} and \vec{A} are functions of p_k and q_k alone. It is to be noted that \vec{A} cannot be a *c*-number. These C.R. can be obtained from the general relations, (valid for finite rotations)

$$\mathbf{T}\mathbf{C} = \mathbf{C}\mathbf{T} \quad \text{or} \quad \mathbf{T}\mathbf{C}\mathbf{T}^{-1} = \mathbf{C} \quad (13.12)$$

and

$$\mathbf{T}\mathbf{A}'_k = \mathbf{A}_k \mathbf{T} \quad \text{or} \quad \mathbf{A}'_k = \mathbf{T}^{-1} \mathbf{A}_k \mathbf{T} \quad (13.13)$$

by specialising to infinitesimal rotations. The first relation represents *invariance* of \mathbf{C} (as in the case of the Hamiltonian) and the second, the *covariance* of \mathbf{A} with respect to rotations. The existence of such a unitary transformation \mathbf{T} follows from the fact that the totality of the \mathbf{A}'_k has the same eigenvalues and satisfies the same C.R. as the \mathbf{A}_k . We can verify that the equations (13.12) and (13.13) follow from eq. (13.7) and the fundamental C.R. (5.21). In particular these relations hold for $\mathbf{A}_k = p_k^{(1)}$ or $\mathbf{A}_k = q_k^{(1)}$. Further (13.10) holds for $\mathbf{C} = \mathbf{P}^2 = \mathbf{P}_1^2 + \mathbf{P}_2^2 + \mathbf{P}_3^2$. Again

$$\mathbf{P}^2 \mathbf{P}_k - \mathbf{P}_k \mathbf{P}^2 = 0 \quad (13.14)$$

follows directly from eq. (13.9). From this we infer that it is possible to bring \mathbf{P}^2 and one of the components \mathbf{P}_k simultaneously into a diagonal form.

It is easy to evaluate, by elementary algebraic methods, all finite Hermitian matrices, which obey the relations (13.9).² If we bring \mathbf{P}^2 and \mathbf{P}_3 into diagonal form, we find the eigenvalues of \mathbf{P}^2 to be given by

$$\mathbf{P}^2 = \hbar^2 j(j+1), \quad (13.15)$$

where j is either a non-negative integer ($j = 0, 1, 2, \dots$) or exceeds it by $\frac{1}{2}$ ($j = \frac{1}{2}, \frac{3}{2}, \dots$). We can combine these into the statement that j can be half-integral. To a given eigenvalue of \mathbf{P}^2 belong $2j+1$ different values of \mathbf{P}_3 ; namely

$$\mathbf{P}_3 = \hbar m \quad \text{with} \quad -j \leq m \leq +j, \quad (13.16)$$

where m changes in steps of unity and is half-integral or integral according as j is half-integral or integral. The matrix elements of \mathbf{P}_1 and \mathbf{P}_2 , for fixed j , are then

$$\left. \begin{aligned} (\mathbf{P}_1 + i\mathbf{P}_2)_{m+1, m} &= \hbar \sqrt{j(j+1) - m(m+1)} = \hbar \sqrt{(j-m)(j+1+m)}, \\ (\mathbf{P}_1 - i\mathbf{P}_2)_{m, m+1} &= (\mathbf{P}_1 + i\mathbf{P}_2)_{m+1, m}; \quad (\mathbf{P}_2)_{m, m} = m \hbar. \end{aligned} \right\} \quad (13.17)$$

All other matrix elements of $(\mathbf{P}_1 - i\mathbf{P}_2)$ and $(\mathbf{P}_1 + i\mathbf{P}_2)$ vanish. For every j the

²Cf., e.g., M. Born and P. Jordan, Elementare Quantenmechanik, Springer Verlag, Berlin, (1930).

matrices (13.17) correspond to an irreducible representation of infinitesimal rotations.

From (13.11) follows for every vector A (which does not depend on vectors with c -number components), and in particular for the co-ordinate matrices³

$$\left. \begin{aligned} (A_1 + i A_2)_{j+1, m+1; j, m} &= -A_{j+1, j}\sqrt{(j+m+2)(j+m+1)}, \\ (A_1 - i A_2)_{j+1, m-1; j, m} &= A_{j+1, j}\sqrt{(j-m+2)(j-m+1)}, \\ (A_3)_{j+1, m; j, m} &= A_{j+1, j}\sqrt{(j+m+1)(j-m+1)}, \end{aligned} \right\} \quad (13.18a)$$

$$\left. \begin{aligned} (A_1 + i A_2)_{j, m+1; j, m} &= A_{j, j}\sqrt{(j+m+1)(j-m)}, \\ (A_1 - i A_2)_{j, m-1; j, m} &= A_{j, j}\sqrt{(j+m)(j-m+1)}, \\ (A_3)_{j, m; j, m} &= A_{j, j}, m. \end{aligned} \right\} \quad (13.18b)$$

$$\left. \begin{aligned} (A_1 + i A_2)_{j-1, m+1; j, m} &= A_{j-1, j}\sqrt{(j-m)(j-m-1)}, \\ (A_1 - i A_2)_{j-1, m-1; j, m} &= -A_{j-1, j}\sqrt{(j+m)(j+m-1)}, \\ (A_3)_{j-1, m; j, m} &= A_{j-1, j}\sqrt{(j+m)(j-m)}. \end{aligned} \right\} \quad (13.18c)$$

The matrix elements vanish for all other pairs of values of (j, m) in the initial and final states. These give the selection and intensity rules for j and m .

Finally we shall add a remark on the composition of two systems with angular momenta j_1 and j_2 . We imagine that the corresponding operators $P_3^{(1)}$ and $P_3^{(2)}$ with the eigenvalues m_1 and m_2 which run over the values $-j_1$ to j_1 and from $-j_2$ to j_2 respectively, are diagonalised simultaneously. We now form the total angular momentum $P_r = P_r^{(1)} + P_r^{(2)}$, and its square $P^2 = \sum_{k=1}^3 P_k^2$. The operator P_3 is already in the diagonal form and the eigenvalues are given by

$$m = m_1 + m_2$$

which occurs as often as there are appropriate numbers in $(-j_1, +j_1)$ and $(-j_2, +j_2)$. If we set $j_1 \geq j_2$, we find the number of m -values $Z(m)$ to be given by:

$$\begin{aligned} j_1 - j_2 \leq m &\leq j_1 + j_2, & Z(m) &= j_1 + j_2 - m + 1, \\ -(j_1 - j_2) \leq m &\leq j_1 - j_2, & Z(m) &= 2j_2 + 1, \\ -(j_1 + j_2) \leq m &\leq -(j_1 - j_2), & Z(m) &= j_1 + j_2 + m. \end{aligned}$$

If we now bring P^2 to diagonal form,⁴ for each m instead of for each m_1 and m_2 separately, then we obtain a series of states with different j such that for each value of j the value of m runs from $-j$ to $+j$. If the value j appears $N(j)$ times, then we obtain the total number of states $Z(m)$ with a fixed m from

$$Z(m) = \sum_{j \geq m} N(j).$$

This is valid for $m \geq 0$, to which cases we can restrict ourselves, since the case $m < 0$

³ For the proof, cf., e.g., M. Born and P. Jordan, Elementare Quantenmechanik, Berlin (1930); P.A.M. Dirac, Quantenmechanik, Leipzig (1930).*

*See also L.D. Landau and E.M. Lifschitz, Quantum Mechanics, Pergamon Press, London (1958).

*This is done for each m through a unitary matrix $S(m_1, j)$. We can calculate it explicitly, cf., e.g., van der Waerden, Die Gruppentheoretische Methode in der Quantenmechanik, Berlin (1932) Sec. 18. See also H.A. Kramers and H.C. Brinkmann, loc. cit. footnote 6 below.

will not yield anything new. Further from this follows that

$$N(j) = Z(j) - Z(j+1).$$

This gives, in our case, that all the j -values differing by steps of unity lie in the interval

$$|j_1 - j_2| \leq j \leq j_1 + j_2 \quad (13.19)$$

and only these are present, each j , in fact, appearing exactly once. This is in agreement with the "vector model" of the old quantum theory. For representations of finite rotations it then follows, as can be seen easily, on using the definition of products of representations introduced earlier, that

$$(D_{j_1}) \times (D_{j_2}) = \sum_{j=|j_1-j_2|}^{j_1+j_2} (D_j). \quad (13.20)$$

Till now we have considered the cases of integral and half-integral j values on the same footing. In Sec. 6, however, it was proved that for a single particle in a central field the angular momentum quantum number, designated there by l instead of j , must always be *integral*. From (13.19) the same result follows for the many-particle case and this cannot be altered even by introducing arbitrary forces between the particles (on account of their continuity). For the necessary generalisation to include spin it is, however, essential that the C.R. (13.9) for the angular momentum matrices as well as the C.R. (13.10) and (13.11) are compatible for half-integral j and m . Only from the definition (13.7) of angular momentum, along with the C.R. for the co-ordinates and momenta follows the integral nature of these quantum numbers.⁵ Therefore, we have still to generalise the definition (13.7) to include the half-integral case.

We shall now give the irreducible representations of the finite rotations (second method) which correspond to the various integral and half-integral values⁶ of j . Here it is useful to start with the group SU_2 of unitary transformations with the determinant 1 of two complex variables ξ_1, ξ_2 instead of with the group of rotations in the three-dimensional space. These transformations have the form

$$\left. \begin{aligned} \xi'_1 &= \alpha^* \xi_1 + \beta^* \xi_2, \\ \xi'_2 &= -\beta \xi_1 + \alpha \xi_2 \end{aligned} \right\} \quad (13.21)$$

with

$$\alpha \alpha^* + \beta \beta^* = 1. \quad (13.22)$$

The corresponding transformation of a_1 and a_2 which leaves the linear form

$$a_1 \xi_1 + a_2 \xi_2$$

invariant is

$$\left. \begin{aligned} a'_1 &= \alpha a_1 + \beta a_2, \\ a'_2 &= -\beta^* a_1 + \alpha^* a_2. \end{aligned} \right\} \quad (13.23)$$

⁵ For a proof of this result using matrix mechanics cf. M. Born and P. Jordan, Elementare Quantenmechanik, p. 164, Springer Verlag, Berlin (1930).

⁶ Besides the textbooks cited earlier on group theory and quantum mechanics cf. for this also H.A. Kramers, Proc. Amsterdam 33, 953 (1930) and H.C. Brinkman, Dissert. Utrecht (1932), where, in particular, applications to the calculation of various matrix elements can be found.

According to (13.21) the $v + 1$ products

$$\xi_1^v, \quad \xi_1^{v-1} \xi_2, \quad \dots, \quad \xi_2^v$$

will transform linearly among themselves. This yields an irreducible representation of degree $v + 1$ of the group considered. If we introduce, as new basis vectors

$$\Xi_r = \frac{\xi_1^{v-r} \xi_2^r}{\sqrt{r!(v-r)!}}, \quad r = 0, 1, \dots, v \quad (13.24)$$

we obtain a unitary representation, since, in addition to $\xi_1 \xi_1^* + \xi_2 \xi_2^*$, $(\xi_1 \xi_1^* + \xi_2 \xi_2^*)^v$ is also invariant and this coincides with $\sum_r \Xi_r \Xi_r^*$. If we set again $v = 2j$ and $r = j - m$,

then the connection with the earlier notations is established. We then obtain a representation D_j of degree $2j + 1$ of the SU_2 group. D_0 is the identity representation; $D_{1/2}$ are the original transformations themselves (fundamental representations).

If we consider D_1 , the connection with the group of rotations in three-dimensional space is obtained. For $j = 1$, $v = 2$, let us transform c_0, c_1, c_2 such that

$$\frac{1}{2}c_0 \xi_1^2 + c_1 \xi_1 \xi_2 + \frac{1}{2}c_2 \xi_2^2$$

remains invariant. Since the transformation has the determinant 1, the determinant

$$\begin{vmatrix} c_1, & c_0 \\ c_2, & c_1 \end{vmatrix} = c_1^2 - c_0 c_2$$

is invariant. If we set

$$x + iy = c_2, \quad x - iy = -c_0, \quad z = c_1, \quad (13.25)$$

then

$$x^2 + y^2 + z^2 = |x + iy||x - iy| + z^2 = c_1^2 - c_0 c_2.$$

Therefore, the representations D_1 in the (x, y, z) -space are ordinary rotations. We can easily show that these x, y, z transform like

$$a_1 a_2^* + a_2 a_1^*, \quad i(a_1 a_2^* - a_2 a_1^*), \quad a_1 a_1^* - a_3 a_3^*,$$

if the a_1 and a_2 transform according to (13.23). Since these are real numbers, we are dealing with real rotations. To a rotation through the Euler angles ϑ, ϕ, ψ correspond the transformation coefficients appearing in eqs. (13.21) and (13.23):

$$\alpha = \cos \frac{\vartheta}{2} e^{\frac{i}{2}(\phi + \psi)}, \quad \beta = i \sin \frac{\vartheta}{2} e^{\frac{i}{2}(\phi - \psi)}. \quad (13.26)$$

For $\vartheta = 0$, we obtain the special case of rotation around the z -axis by an angle $\phi + \psi$, for which the matrix will be diagonal. The identity transformation for spatial rotations corresponds not only to the identity belonging to transformations (13.21) but also to the transformation

$$\xi'_1 = -\xi_1, \quad \xi'_2 = -\xi_2.$$

Hence the rotations are abbreviated representations of SU_2 ; conversely $D_{1/2}$ is a two-valued representation of the rotation group. With the help of (13.26) we can also calculate taking into account the quantities Ξ_r , which are defined in (13.24), the matrices of the general representation D_j as functions of the angles ϑ, ϕ, ψ .⁷ They are already reduced with respect to the sub-group of rotations around the z -axis.

⁷ P. Güttinger, Z. Physik 73, 169 (1931).

We obtain two-valued representations for half-integral j and single-valued representations for integral j . The spherical functions of the l^{th} order transform like D_l .

We must further mention that the Hamiltonian is also invariant under the inversion

$$x'_k = -x_k \quad (13.27)$$

about the centre of mass of the system. This inversion commutes with all rotations. Thus, on inversion, every eigenfunction will be multiplied simply by a factor. (We can imagine the degeneracy arising from the rotation group to be removed completely by an external magnetic field without thereby destroying the invariance of the Hamiltonian with respect to this inversion.) Since two successive inversions give the identity transformation, the factor (ϵ), which is defined by

$$\psi(-q_1, \dots, -q_r) = \epsilon \psi(q_1, \dots, q_r) \quad (13.28)$$

can be equal to only ± 1 :

$$\epsilon = \pm 1. \quad (13.29)$$

We shall call this factor the *parity* of the state. A wave function is said to be of even or odd parity, according as $\epsilon = +1$ or $\epsilon = -1$. For a single particle in a central field it follows from the properties of the spherical functions that

$$\epsilon = (-1)^l.$$

Therefore, in general, for many particles we have

$$\epsilon = (-1)^{l_1+l_2+\dots+l_r}. \quad (13.30)$$

This holds only for particles that are not interacting, but interaction cannot alter it because it is switched on continuously. The matrices of the co-ordinates of the particles are different from zero only if ϵ has different signs in the initial and final states (Laporte's Rule).

(b) Wave Functions for Particles with Spin

With these tools, it is now easy for us to set up the *generalised wave equations for particles with spin*. Originally,⁸ one had only the electrons and protons in view. We shall, however, understand by the spin of a particle an angular momentum in general, which cannot be traced to the translational motion of the particle and the magnitude of which (in contrast to those of its components) will be considered as a fixed number.

The description of spin rests on the following assumptions:

1. Besides the orbital angular momentum

$$\mathbf{l}_1 = \frac{i}{\hbar} \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \right) \quad (13.31)$$

there is a spin angular momentum with the components s_1, s_2, s_3 . Operators which act on the wave functions are associated with them. These operators commute with the position and momentum co-ordinates of the particle. Here we consider both the (spin and orbital) angular momenta to be measured in units of \hbar .

⁸ W. Pauli, Z. Physik 43, 601 (1927).

2. The application of the operators

$$\omega_k = -i(l_k + s_k) \quad (13.32)$$

to the eigenfunctions corresponds to infinitesimal rotations, so that

$$P_k = \hbar(l_k + s_k) \quad (13.33)$$

plays the role of the total angular momentum. In fact the operator P_k commutes with all operators invariant under rotation and is a constant of motion if the Hamiltonian is invariant under rotation. If the Hamiltonian is invariant only under rotations about an axis, say the x_3 -axis, then P_3 is always a constant. From this assumption follows, purely kinematically, the C.R. (13.6) for the ω_k and, therefore, since the l_k commute with the s_k , the latter obey the C.R. similar to (13.9):⁹

$$s_1 s_2 - s_2 s_1 = i s_3 \dots \quad (13.34)$$

Since $s_1^2 + s_2^2 + s_3^2$ is a fixed number we have from (13.34)

$$s_1^2 + s_2^2 + s_3^2 = s(s+1) \cdot 1, \quad (13.35)$$

with integral or half-integral values for s . We can introduce in the wave function, in addition to the position co-ordinates x_k of the particle, also one of the components s_k , say s_3 as an independent variable, so that the wave function is of the form

$$\psi(x, s_3; t). \quad (13.36)$$

But s_3 is, as we have seen, capable of taking only the values $-s, -(s-1), \dots, +s$. We can, therefore, write instead of (13.36)

$$\psi(x, s_3; t) = \sum_{\mu} C_{\mu}(s_3) \psi_{\mu}(x, t). \quad (13.36')$$

where the $C_{\mu}(s_3)$ which do not depend on x and t are defined by

$$C_{\mu}(s_3) = \begin{cases} 1 & \text{for } s_3 = \mu, \\ 0 & \text{otherwise.} \end{cases} \quad (13.37)$$

The $C_{\mu}(s_3)$ are normalised and orthogonal, i.e. we have

$$\sum_{s_3=-s}^{+s} C_{\mu}^*(s_3) C_{\mu'}(s_3) = \begin{cases} 1 & \text{for } \mu = \mu', \\ 0 & \text{for } \mu \neq \mu'. \end{cases} \quad (13.37')$$

For applications, the particular form of the $C_{\mu}(s_3)$ in eq. (13.37) is not essential but only the relations (13.37') which remain unchanged under rotations of the co-ordinate axes are essential.

In general $\psi_{\mu}^* \psi_{\mu}$ is again the probability in position-spin-space. The density ρ in the space of the position co-ordinates alone is, therefore,

$$\rho = \sum_{\mu} \psi_{\mu}^*(x, t) \psi_{\mu}(x, t) = \sum_{\mu} \psi_{\mu}^*(x, t) \psi_{\mu}(x, t). \quad (13.38)$$

Its volume integral is constant in time and hence the orthogonality and normalisation

⁹ Originally, this C.R. was based on the analogy to the C.R. for l_k ; the latter can be deduced from the canonical C.R. for p_k and q_k . J.v. Neumann and E. Wigner, Z. Physik 47, 203 (1927) were the first to indicate the possibility of the kinematical derivation of the C.R. for the s_k from the rotation group.

conditions for the eigenfunctions belonging to the stationary states are

$$\psi(x, s_3) = \sum_{\mu} C_{\mu}(s_3) \psi_{\mu}(x),$$

$$\sum_{s_3} \int dx \psi_n(x, s_3) \psi_m(x, s_3) = \sum_{\mu} \int dx \psi_{n\mu}^*(x) \psi_{m\mu}(x) dx = \begin{cases} 1 & \text{for } m = n, \\ 0 & \text{for } m \neq n. \end{cases} \quad (13.38')$$

We shall speak about the current operators when we come to the relativistic theory.

The way in which the operators s_1, s_2, s_3 act on the wave functions is seen most easily from their matrix representation (similar to (13.17)) which reads

$$\left. \begin{aligned} (s_1 + i s_2)_{\mu+1,\mu} &= \sqrt{(s - \mu)(s + 1 + \mu)}, \\ (s_1 - i s_2)_{\mu-1,\mu} &= \sqrt{(s - \mu + 1)(s + \mu)}, \\ (s_3)_{\mu\mu} &= \mu. \end{aligned} \right\} \quad (13.39)$$

We have then, in general

$$(s_k) \psi_{\mu}(x, t) = \sum_{\mu'} \psi_{\mu'} (s_k)_{\mu',\mu}, \quad (13.40)$$

where, however, at the most, only two of the matrix elements $(s_k)_{\mu'\mu}$ are different from zero. For example

$$\begin{aligned} s_1 \psi_{\mu} &= \frac{1}{2} (s_1 + i s_2) \psi_{\mu} + \frac{1}{2} (s_1 - i s_2) \psi_{\mu} \\ &= \frac{1}{2} \psi_{\mu-1} \sqrt{(s - \mu + 1)(s + \mu)} + \frac{1}{2} \psi_{\mu+1} \sqrt{(s - \mu)(s + 1 + \mu)}, \end{aligned}$$

where for $\mu = s$, the second and for $\mu = -s$, the first term on the right-hand side vanishes. On the other hand, we have simply

$$s_3 \psi_{\mu} = \psi_{\mu} \mu.$$

The Hamiltonian will contain in general s_k besides x . For example, it will contain the additional term $K(s_1 \mathcal{H}_1 + s_2 \mathcal{H}_2 + s_3 \mathcal{H}_3)$ in an external magnetic field with components $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$, where the numerical factor K denotes the ratio between the magnetic moment and the angular momentum. In general the form of the Hamiltonian must be determined by comparison with the classical theory.

Under spatial rotations the C_{μ} and ψ_{μ} in eq. (13.36') transform like the coefficients c_{μ} and the variables Ξ_{μ} of the invariant form

$$\sum c_{\mu} \Xi_{\mu},$$

where, according to eq. (13.24), with $v = 2s$, $r = s - \mu$,

$$\Xi_{\mu} = \frac{\xi_1^{\mu} + \mu \xi_2^{\mu}}{\sqrt{(s + \mu)(s - \mu)}}$$

and ξ_1, ξ_2 transform according to eq. (13.21). The transformation of ψ_{μ} will, therefore, be determined directly by the representation (D_s) of the rotation group which is single- or double-valued according as s is half-integral or integral. Under the inversion (13.27) in the origin under which the s_k (which are, properly speaking, anti-symmetric tensors) are invariant, the indices of ψ_{μ} can also be left unaltered. For $s = 1$, the wave functions $\psi_1, \psi_0, \psi_{-1}$, transform under the rotation D^{-1} just like $-(x_1 - ix_2), x_3, x_1 + ix_2$ respectively do under the rotation D , so that suitable linear combinations of ψ_{μ} can be interpreted as vector components.

Of special interest, however, is the case, $s = \frac{1}{2}$, because this is the spin of the elementary particles (like the electron and the nucleon)* as experiments show. For this case, according to eqs. (13.39)

$$(s_1 + i s_2)_{+\frac{1}{2}, -\frac{1}{2}} = (s_1 - i s_2)_{-\frac{1}{2}, +\frac{1}{2}} = 1,$$

and, therefore, in the matrix form

$$s_1 + i s_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad s_1 - i s_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad s_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix},$$

or with

$$s_k = \frac{1}{2} \sigma_k, \quad (13.41)$$

we have

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (13.42)^{**}$$

If we write $\psi(\vec{x}; t)$ as a column vector:

$$\psi(\vec{x}; t) = \begin{pmatrix} \psi_1(\vec{x}; t) \\ \psi_2(\vec{x}; t) \end{pmatrix} \quad (13.43)$$

(and, correspondingly, ψ^* as a row vector), then we can define the "density components of the spin angular momentum":

$$d_k = (\psi^* \sigma_k \psi).$$

With the help of (13.42) we then obtain:

$$\left. \begin{aligned} d_1 &= \psi_2^* \psi_1 + \psi_1^* \psi_2, \\ d_2 &= i(\psi_2^* \psi_1 - \psi_1^* \psi_2), \\ d_3 &= \psi_1^* \psi_1 - \psi_2^* \psi_2. \end{aligned} \right\} \quad (13.44)$$

These quantities transform like the components of a vector

$$\vec{d} = (\psi^* \vec{\sigma} \psi),$$

while

$$\varrho = \psi_1^* \psi_1 + \psi_2^* \psi_2 \quad (13.45)$$

is invariant.

The matrices σ_k satisfy the relations

$$\left. \begin{aligned} \sigma_1 \sigma_2 &= -\sigma_2 \sigma_1 = i \sigma_3, \dots, \\ \sigma_1^2 &= \sigma_2^2 = \sigma_3^2 = 1. \end{aligned} \right\} \quad (13.46)$$

They are more restrictive than the relations (13.34) and (13.35) and signify that the σ_k multiply among themselves according to the multiplication rule for the units of the quaternions multiplied by i .

Under spatial rotations, ψ_1 , ψ_2 and $-\psi_2^*$, ψ_1^* transform like a_1 , a_2 in (13.23) and, therefore, according to the matrix equation

$$\begin{pmatrix} \psi'_1 & -\psi'^*_2 \\ \psi'_2 & \psi'^*_1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \begin{pmatrix} \psi_1 & -\psi_2^* \\ \psi_2 & \psi_1^* \end{pmatrix}.$$

* Particles with half-integral spin are called fermions, integral-spin particles being called bosons.

** These are the famous Pauli matrices.

Here α and β are related to the angles of rotation through (13.26). As already mentioned,

$$\omega_k = -i s_k = -\frac{i}{2} \sigma_k$$

then determines the transformation of ψ_1 and ψ_2 for infinitesimal rotations.

The case of more than one particle, say N , with spin which interact with one another can be settled without further ado. We have to introduce the eigenfunctions

$$\psi(x_{r1}, x_{r2}, x_{r3}, s_{rk}) \quad (13.47)$$

where the index r labels the particles and runs from 1 to N . Each s_{rk} runs from $-s_r$ to $+s_r$. Written with indices, the eigenfunctions can be put in the form

$$\psi(x_{rk}, s_{rk}, t) = \sum_{\mu_1, \dots, \mu_N} C_{\mu_1}(s_{13}) \dots C_{\mu_N}(s_{N3}) \psi_{\mu_1, \dots, \mu_N}(x_{11} \dots x_{N3}), \quad (13.48)$$

where μ_r runs from $-s_r$ to $+s_r$. In the case of electrons each μ can take only two values $+\frac{1}{2}$ and $-\frac{1}{2}$ and 2^N functions are defined by the eq. (13.48). An operator s_{rk} acts only on the index μ_r with the same label r as in the single particle case.

An essential difference between composite structures (like the deuteron and helium) and elementary particles* (like the leptons and nucleons) as well as a basis for the value $\frac{1}{2}$ for the spin of the fermions is to be found only in the relativistic quantum mechanics (cf. Sec. 18).

14. The Behaviour of Eigenfunctions of Many Identical Particles Under Permutation.¹ The Exclusion Principle

(a) Permutations and the Symmetry of Eigenfunctions

When we deal with several identical particles, many special properties arise, which depend on the fact that the Hamiltonian is always invariant under any permutation of the particles. If the particles possess spin, the spin co-ordinates s_{rk} must also be permuted simultaneously with the position co-ordinates x_{rk} ($k = 1, 2, 3$). If the

*Elementary particles can be classified as: the photons, the leptons and the hadrons. The electrons, the mu-meson and the neutrino constitute the class of leptons. Nucleon is a collective name for the proton and the neutron both of which belong to the hadron family. For details regarding particle properties, reference may be made to "Review of Particle Properties" by Particle Data Group, Review of Modern Physics, April (1978).

¹Cf. the text-books on Group Theory cited earlier. Historically speaking, the following remarks may be made. The problem of many identical particles was first treated quantum-mechanically by P.A.M. Dirac, Proc. Roy. Soc. Lond., Ser. A 112, 661 (1926) (but without spin) and W. Heisenberg, Z. Physik 40, 501 (1926) (in this paper is to be found for the first time the important application to the He-Spectrum including spin); in both of these papers is also to be found the general quantum-mechanical formulation of the exclusion principle (W. Pauli, Z. Physik 31 765 (1925)). The statistics of particles with symmetric states was first established by S.N. Bose, Z. Physik 26, 178 (1924) and A. Einstein (Berl. Ber. 1924, 261; 1925, 1) and that of particles with antisymmetric states by E. Fermi, Z. Physik 36, 902 (1926) and P.A.M. Dirac (I.e.). The general case of N particles and its connection with group theory is to be found for the first time in E. Wigner, Z. Physik 40, 883 (1927). The proof of the fact that the protons as well as the electrons have spin $\frac{1}{2}$ and they obey the exclusion principle was given

Hamiltonian contains the position variables alone, then there is invariance even under the permutation of the position co-ordinates. If the spin-dependent part of the Hamiltonian is relatively small, then the invariance under permutation of position holds approximately. We shall come back to this point later. At present we consider the simultaneous permutation of the spin and position variables under which *exact* invariance holds. If, therefore P is a permutation of the N numbers $1, 2, \dots, r, \dots, N$, which label N identical particles, then we obtain from each eigenfunction $\psi(x_{11} \dots x_{N3}, s_{13} \dots s_{N3}, t)$, by applying the permutation P , a new eigenfunction, which belongs to the same observable state of the system:

$$\psi'(x_{11} \dots s_{N3}, t) = P\psi(x_{11} \dots s_{N3}, t).$$

In fact every operator symmetric in the particle variables, and in particular, the energy operator has the same expectation value if we make calculations with ψ' or ψ . *Only these symmetric operators, however, correspond to observable quantities in the case of identical particles.* On account of the indistinguishability of one particle from another, it is only meaningful to ask, e.g. for the probability that *one* of the particles has the position x_{1k} and the spin s_{13} and *another* the position x_{2k} and the spin s_{23} , etc., but not the probability that the *first* particle has the position and the spin x_{1k}, s_{13} and the *second*, the position and spin x_{2k}, s_{13} . The probability for the first case is given by

$$\sum_P |P|\psi(x_{11} \dots s_{N3})^* dx_{11} \dots dx_{N3}, \quad (14.1)$$

if we always consider the position co-ordinates x_{rk} to be determined only up to an interval dx_{rk} and the spin co-ordinates s_{r3} take on the values between $-s$ and $+s$ (for electrons, $s = \frac{1}{2}$).

From the general results of the previous section it follows that the stationary states of the complete system must be decomposed into sub-systems which correspond to the various irreducible representations of the permutation group. Further, for a symmetric quantity, only those matrix elements are different from zero for which the initial and final states belong to the same term system. If the representation is of order 1, then the levels are not degenerate (accidental degeneracies or those which arise from the invariance of the Hamiltonian under groups other than the permutation group, remain outside the scope of our present consideration); the eigenfunction is multiplied by a numerical factor as a consequence of every permutation. If, in general, the representation is of order h , then the associated energy value is h -fold degenerate. In the corresponding h -dimensional linear vector space of the eigenfunctions, we can find a basis of mutually orthogonal and normalised eigenfunctions u_1, u_2, \dots, u_h which go over, by the application of the permutation P , into the new eigenfunctions

$$Pu_s = \sum_{r=1}^h u_r c_{rs}(P) \quad (14.2)$$

by D.M. Dennison, Proc. Roy. Soc. Lond. Ser. A 115, 483 (1927) by an explanation of the drop in the heat of rotation of hydrogen. It was shown by N.F. Mott, Proc. Roy. Soc. Lond., Ser. A 125, 222 (1929) and J.R. Oppenheimer, Phys. Rev. 32, 361 (1928) that the symmetry class of the eigenfunctions is essential in collision problems. The calculation of N.F. Mott, Proc. Roy. Lond., Ser. A 126, 259 (1929) on the collision of two equal point charges then gave, among other things, the result that the He-Nucleus (α -particle) has symmetric states.

corresponding to representation $c(P)$. Since

$$\sum_s \int u_s^* u_s dx = \sum_s \int (P u_s)^* (P u_s) dx,$$

(here each of the s_{rj} is to be summed from $-s$ to $+s$), the $P u_s$ are also orthogonal and normalised, if the u_s are orthonormal and the representation is unitary. From an arbitrary function $f(q_1, \dots, q_N)$ (where q_r is a compact notation for x_{1r}, x_{2r}, x_{3r} , and s_{3r}) we obtain a special function $v(q_1 \dots q_N)$ which transforms as per the representation (14.2), by the construction

$$v(q_1 \dots q_N) = \sum_P A_P \cdot P f(q_1 \dots q_N), \quad (14.3)$$

with a suitable choice of the numerical coefficients A_P .

Special representations of degree 1 belong to the symmetric and antisymmetric functions. In the first case

$$P u(q_1 \dots q_N) = u(q_1 \dots q_N), \quad (14.4)$$

we have the identity representation, i.e. to each group element corresponds the identity element. For the antisymmetric representation we have to differentiate between even and odd permutations. If $\delta_P = 1$ for even and $\delta_P = -1$ for odd functions, we have for an antisymmetric function

$$P u(q_1 \dots q_N) = \delta_P \cdot u(q_1 \dots q_N). \quad (14.5)$$

Since

$$\delta_{PQ} = \delta_P \cdot \delta_Q, \quad \delta_{P^{-1}} = \delta_P, \quad \delta_1 = 1,$$

this is, in fact, a representation and for the validity of (14.3) it is sufficient that u changes sign on the interchange of any two variables. We obtain, from an arbitrary function f , a symmetric function, according to eq. (14.3), if we set all the coefficients A_P equal to 1:

$$v_{\text{symm}}(q_1 \dots q_N) = \sum_P P f(q_1 \dots q_N) \quad (14.6a)$$

and an antisymmetric function, if we set $A_P = \delta_P$ (i.e. $A_P = \pm 1$):

$$v_{\text{antis}}(q_1 \dots q_N) = \sum_P \delta_P P f(q_1 \dots q_N). \quad (14.6b)$$

If only two particles are present, the symmetric and the antisymmetric representations are the only irreducible representations; the energy values, therefore, simply split up into the two states.

If a system of N particles is once in a particular term system (belonging to a certain irreducible representation $c(P)$ of the permutation group), then it cannot be brought into any other system by an external agency (force field, radiation), because the perturbation energy is symmetric in the particles and its matrix elements between initial states of the system considered and the final states of another system vanish. Further, as a consequence of the wave equation the symmetry character of time-dependent wave functions which was present at time zero will remain the same for ever. We, therefore, speak also of *non-combining term systems*. Here, however, the case in which the number, N , of the particles of the kind considered does not remain constant – e.g., the system collides with an additional particle of the kind considered – requires special consideration.

Let, for example, an atomic system with N electrons be given and let us assume it

to be in a state with an eigenfunction $u_0(q_1 \dots q_N)$ which belongs to a definite irreducible representation $D^{[N]}$ of the group \sum_N of permutations of N elements. Then let the $(N + 1)^{\text{th}}$ electron collide with the atom and let the eigenfunctions $U_p(q_1 \dots q_N, q_{N+1})$ of the total system before the collision be so chosen that they belong to a definite irreducible representation $D^{[N+1]}$ of the group of permutations of $(N + 1)$ elements. Then U_p has the form

$$U_p = \sum_{\mathbf{P}} A_{\mathbf{P}, p} u_1(q_1 \dots q_N) v(q_{N+1}),$$

where $v(q_{N+1})$ is the eigenfunction of the colliding electron and the $A_{\mathbf{P}, p}$ are suitable numerical coefficients. The representation $D^{[N]}$ of u must be contained in the representation $D^{[N+1]}$ of U by the reduction with respect to the sub-group \sum_N of \sum_{N+1} . Since, further, the u vanishes rapidly at a large distance R of one of the electrons, we get with a high degree of approximation:

$$U_p(q_1 \dots q_N, R) = A_{1p} u_1(q_1 \dots q_N) v(R).$$

After the collision we obtain a new function of a similar nature:

$$U'_p(q_1 \dots q_{N+1}) = \sum_{\mathbf{P}} A'_{\mathbf{P}, p} P u'_1(q_1 \dots q_N) v'(q_{N+1}).$$

Then U'_p will necessarily belong to the same representation $D^{[N+1]}$ of \sum_{N+1} to which U_p belongs. On the other hand, on applying the permutations P to \sum_N the u' transform according to a (possibly reducible) representation which may contain any irreducible representations $D^{[N]}$ which arise from the reduction of $D^{[N+1]}$ with respect to the sub-group \sum_N of \sum_{N+1} . These are, in general, many representations and the atom can then go from a state of one system to a state of another system by collision with an additional electron. Only in the two special cases, already mentioned, a multiplicity of $D^{[N]}$ does not occur. Namely, if we have to deal with a symmetric or an antisymmetric function $U(q_1 \dots q_{N+1})$ of the co-ordinates of $N + 1$ particles, then the \bar{u}_s and \bar{u}_a which emerge from the decomposition

$$U_s(q_1 \dots q_{N+1}) = \sum_{\mathbf{P}} \mathbf{P} u(q_1 \dots q_N) v(q_{N+1}) = \sum_{k=1}^{N+1} T_{N+1, k} \bar{u}_s(q_1 \dots q_N) v(q_{N+1}),$$

$$U_a(q_1 \dots q_{N+1}) = \sum_{\mathbf{P}} \delta_{\mathbf{P}} \mathbf{P} u(q_1 \dots q_N) v(q_{N+1}) = \sum_{k=1}^{N+1} \delta_k T_{N+1, k} \bar{u}_a(q_1 \dots q_N) v(q_{N+1})$$

must necessarily again be respectively symmetric and antisymmetric in the N variables $q_1 \dots q_N$. Here $T_{N+1, k}$ denotes the interchange (transposition) of two indices $N + 1$ and k and hence $T_{N+1, N+1}$ denotes the identity transformation; $\delta_k = +1$ for $k = N + 1$ and $\delta_k = -1$ for $1 \leq k \leq N$.

(b) The Exclusion Principle

It has been shown experimentally that if we interchange the spin and position variables simultaneously then for each kind of particles only one class of states is present. This class can then be only either a symmetric or an antisymmetric class. Experience shows further that for the electrons and nucleons it is the antisymmetric class alone which is present in nature. For other particles, e.g., the He-nuclei (α -

particles) the symmetric class is present in nature. The fact that quantum mechanics yields *more* states than actually occur in nature (and all of them equally possible), is still a puzzle and it is to be hoped that a future theory of elementary particles will bring a deeper insight into the essence of this restricted choice of nature.²

The properties of the symmetry classes stand out clearly, if we consider particles which, in the first approximation, are not coupled, i.e. are free from any mutual interactions, but may be subject to an external force field. Let, in this case, $u_1(q), \dots u_N(q)$ be the eigenfunctions of the states in which the N electrons are found. Among these, identical states can also be present. Then the symmetric eigenfunction is the sum of the products

$$U_s(q_1 \dots q_N) = \sum_P P u_1(q_1) \dots u_N(q_N), \quad (14.7)$$

where the permutations are over the labels of the particle co-ordinates for fixed indices $1 \dots N$ of the states. (For normalising U_s , a suitable numerical factor is to be added.) Similarly for the antisymmetric eigenfunction

$$U_a(q_1 \dots q_N) = \sum_P \delta_P P u_1(q_1) \dots u_N(q_N) = \begin{vmatrix} u_1(q_1) & \dots & u_N(q_1) \\ u_1(q_2) & \dots & u_N(q_2) \\ \dots & \dots & \dots \\ u_1(q_N) & \dots & u_N(q_N) \end{vmatrix}. \quad (14.8)$$

Here the wave function has been written as a determinant. *The antisymmetric eigenfunction vanishes identically, if two of the states coincide ($u_i(q) = u_k(q)$ for $k \neq i$).* For particles with antisymmetric states it can, therefore, never happen that two particles find themselves in the same state. This is the content of the *exclusion principle* which was formulated already before the wave mechanics of the electrons was set up.* That it is valid for nucleons also, was recognised only later. From the validity of the exclusion principle for a particular kind of particles it follows, conversely, that the particles have antisymmetric states. For, *only* the antisymmetric eigenfunctions have the property that they vanish, if two particles are in the same state.

For the consistent exclusion of all symmetry classes except one, for a given type of particles, it is essential that the nature of symmetry class itself cannot be observed within the domain of validity of classical mechanics (geometrical optics). If we consider for the sake of simplicity only two particles, we can differentiate between them, e.g., by using the continuity in the changes in their positions *only if their wave-functions $\psi_1(q, t)$ and $\psi_2(q, t)$ never overlap, i.e. are different from zero in disjoint spatial domains*. (Strictly speaking, separate regions in position-spin space suffice; in the following, we write $\int dq$ instead of $\sum_{S_3} \int dx_1 dx_2 dx_3$.) In this case

$$\psi_1^*(q_r, t) \psi_2(q_r, t) = 0 \quad r = 1, 2$$

² Several attempts have been made to enforce these restrictions on the possible types of particles by introducing suitable singularities in the interaction energy of two elementary particles, when their positions and spins coincide. One should then be able to achieve that only the antisymmetric eigenfunctions remain regular. This was done in a mathematically correct manner by G. Jaffé, Z. Physik **66**, 748 (1930). The singularities are, however, of such a kind that they can hardly correspond to reality.

* W. Pauli, loc. cit. in footnote 1.

in the whole q -space as well as in all the time intervals considered. Therefore, for the expectation value of an arbitrary operator $F(p_1, p_2, q_1, q_2)$ which is symmetric in both the particles, both in the symmetric and antisymmetric cases, and with the normalised eigenfunctions

$$\left. \begin{aligned} \Psi_s(q_1, q_2, t) &= \frac{1}{\sqrt{2}} [\psi_1(q_1, t) \psi_2(q_2, t) + \psi_1(q_2, t) \psi_2(q_1, t)], \\ \Psi_a(q_1, q_2, t) &= \frac{1}{\sqrt{2}} [\psi_1(q_1, t) \psi_2(q_2, t) - \psi_1(q_2, t) \psi_2(q_1, t)], \end{aligned} \right\} \quad (14.9)$$

we have

$$\left. \begin{aligned} \bar{F}(t) &= \int \Psi_s^* F \Psi_s d q_1 d q_2 = \int \Psi_a^* F \Psi_a d q_1 d q_2, \\ &= \int \psi_1^*(q_1, t) \psi_2^*(q_2, t) [F \psi_1(q_1, t) \psi_2(q_2, t)] d q_1 d q_2 \\ &= \int \psi_1^*(q_1, t) \psi_2^*(q_1, t) [F \psi_1(q_2, t) \psi_2(q_1, t)] d q_1 d q_2, \end{aligned} \right\} \quad (14.10)$$

since in this case

$$\int \psi_1^*(q_1, t) \psi_2^*(q_2, t) [F \psi_1(q_2, t) \psi_2(q_1, t)] d q_1 d q_2 = 0$$

(at least if F is a rational integral function of p). We can restate the fact that the electrons satisfy the exclusion principle and have antisymmetric states by saying that all electrons must "conclude a pact with one another" or "know of one another" in order to satisfy this principle. We see, however, that this "pact" automatically comes into effect only if the wave-packets of the electrons overlap; i.e. if the possibility that both are at the same point in the position-spin space is not excluded in principle (without consideration of the symmetry class).

We shall now consider the behaviour of several electrons somewhat more precisely, with regard to the separation of the position and spin co-ordinates. In many cases, it is permissible to consider the spin orbit interaction, (i.e. the part of the Hamiltonian containing the spin operators of the electrons) as small compared to the mutual interaction of the electrons. In the zeroth approximation (each electron moves in the same external force field) the Hamiltonian has the form

$$H^{(0)} = \sum_{r=1}^N H_r^{(0)},$$

where each $H_r^{(0)}$ acts only on the position co-ordinates of the r^{th} electron. In the first approximation a perturbation

$$H^{(1)} = \sum_{r,s} V_{rs} \frac{e^2}{r_s}$$

is added, which is symmetric in the position co-ordinates of the particles. This is the Coulomb interaction energy of the particles introduced in eqs. (5.12) and (5.13). In the second approximation the interaction energy between spin and orbital motion enters for the first time:

$$H^{(2)} = V(x_{rs}, s_{rs}).$$

If $H^{(2)}$ can be considered not only smaller than $H^{(0)}$ but also smaller than $H^{(1)}$, we speak of the Russell-Saunders coupling.

The behaviour of the eigenfunctions based on these assumptions can now be in-

vestigated for the simple case of *two* particles. According to eqs. (13.36') and (13.43) together with

$$\left. \begin{aligned} u(x, s_3) &= u(x) [a_\alpha C_+(s_3) + a_\beta C_-(s_3)], \\ v(x, s_3) &= v(x) [b_\alpha C_+(s_3) + b_\beta C_-(s_3)], \end{aligned} \right\} \quad (14.11)$$

the solutions which are totally antisymmetric in the combined spin and space co-ordinates are, in the zeroth and the first approximations, of the form

$$U(x_1, x_2, s_{13}, s_{23}) = u(x_1, s_{13}) v(x_2, s_{23}) - u(x_2, s_{23}) v(x_1, s_{13}). \quad (14.12)$$

In the above, we drop the index k ($= 1, 2, 3$) of the three spatial co-ordinates of each particle. Equation (14.12) states that in the zeroth and the first approximations the spin and position variables are separable; each of the functions $u(x)$ and $v(x)$ is the space eigenfunction of one of the electrons in *one* of the two states considered. If both electrons are in the same state, in regard to the position co-ordinates, then we have to set $u(x) = v(x)$. If originally no external field was present and the Hamiltonian is consequently invariant under rotation, then we obtain for an *arbitrary* choice of $a_\alpha, a_\beta, b_\alpha, b_\beta$ admissible eigenfunctions. Then $U(x_1, x_2, s_{13}, s_{23})$ can be expressed as a linear combination of the following four orthonormal eigenfunctions:

$$U^1(x_1, x_2, s_{13}, s_{23}) = \frac{1}{\sqrt{2}} [u(x_1) v(x_2) - u(x_2) v(x_1)] A_{m_s}(s_{13}, s_{23}) \quad (14.13a)$$

with $m_s = -1, 0, +1$ and

$$\left. \begin{aligned} A_{-1}(s_{13}, s_{23}) &= C_-(s_{13}) C_-(s_{23}), \\ A_0 &= \frac{1}{\sqrt{2}} [C_+(s_{13}) C_-(s_{23}) + C_-(s_{13}) C_+(s_{23})], \\ A_{+1} &= C_+(s_{13}) C_+(s_{23}), \end{aligned} \right\} \quad (14.13b)$$

and

$$\left. \begin{aligned} U^{II}(x_1, x_2, s_{13}, s_{23}) &= \frac{1}{\sqrt{2}} [u(x_1) v(x_3) + v(x_1) u(x_3)] \times \\ &\quad \times \frac{1}{\sqrt{2}} [C_+(s_{13}) C_-(s_{23}) - C_-(s_{13}) C_+(s_{23})]. \end{aligned} \right\} \quad (14.14)$$

We see that the spin eigenfunctions, appearing as factors here, are symmetric in the first case and antisymmetric in the second case. In the latter case, any operator $s_k = s_{1k} + s_{2k}$ acting on the spin eigenfunction gives the value zero. From this itself it follows that the spin eigenfunction is invariant under rotations. This can also be seen directly since it is multiplied by the determinant for any linear transformation which is carried out in the same manner on $C_+(s_{13})C_-(s_{13})$ and $C_+(s_{23})C_-(s_{23})$. This determinant has, however, as we have seen, the value 1 for the transformations of C_+ and C_- associated with rotations. To U^1 , therefore, belongs a term with $S=0$ (singlet term). Even, when, on account of the invariance of the Hamiltonian under rotations, several eigenfunctions u and v are degenerate³ (non-vanishing resultant angular momentum L of the orbital motion) there does not occur a further splitting of the terms on introducing the perturbation energy $H^{[2]}$. Note that if $u(x) = v(x)$, the first normalisation factor $1/\sqrt{2}$ in U^1 has to be replaced by $\frac{1}{2}$. In the

³ Then suitable linear combination of the form $u_\alpha(x_1) v_\lambda(x_2) + u_\lambda(x_2) v_\alpha(x_1)$ must be constructed.

first approximation, i.e. neglecting $H^{[2]}$, there belongs to U^1 the perturbation term

$$\Delta E_{11} = J_0 + J_1, \quad (14.15 \text{ a})$$

where

$$\left. \begin{aligned} J_0 &= \int |u(x_1)|^2 |v(x_2)|^2 V(x_1, x_2) dx_1 dx_2, \\ J_1 &= \int u^*(x_1) u(x_2) v^*(x_1) v(x_2) V(x_1, x_2) dx_1 dx_2. \end{aligned} \right\} \quad (14.16)$$

Here J_1 is the exchange integral.

In the first case, which corresponds to the eigenfunctions U^1 , the spin eigenfunctions are symmetric. The A_{-1}, A_0, A_1 transform into one another under rotations. This is due to the fact that the rotations and the permutations commute; the symmetry character of the eigenfunctions will not, therefore, be altered by rotations. The spin eigenfunctions A_{-1}, A_0, A_1 correspond respectively to the values $-1, 0, 1$ of the quantum number m_s of the third component $s_3 = s_{13} + s_{23}$ of the resultant spin angular momentum and to the value $s = 1$ of the quantum number corresponding to the magnitude of the resultant spin momentum. In the first approximation, the associated perturbation energy is,

$$\Delta E_1 = J_0 - J_1. \quad (14.15 \text{ b})$$

if J_0 and J_1 denote the same integrals as in eq. (14.16). If L is the orbital angular momentum, the perturbation $H^{[2]}$ introduces a further splitting of the terms in general, on account of the fact that the reducible representation

$$D_1 \times D_2$$

of the rotation group decomposes into its irreducible components

$$D_J$$

with $J = L + 1, L, L - 1$. For $L = 0$ (S -state), evidently no splitting occurs. Here, therefore, we are dealing with a triplet state. It is to be remembered, that for $u(x) = v(x)$, the eigenfunction U^1 vanishes identically.

The following is an important result: *The exclusion principle gives rise, for two electrons which belong to states which are symmetric in the position co-ordinates alone, to a singlet state and for those that belong to states which are antisymmetric in the position co-ordinates, to triplet states. Even when the interaction between spin and the position co-ordinates is neglected, these states differ in energy by the exchange integral.* For the states, not present in nature* which belong to the totally symmetric class, the association of the multiplicity with the symmetry class in the position co-ordinates alone would be just the opposite.

This is essentially the content of Heisenberg's theory of the Helium spectrum. On account of the indistinguishability of electrons, the interchange of positions of two electrons is, in principle, never observable. By the "exchange" of the electrons in a Helium atom, we can, at most observe that the spins of the outer and of the inner electrons, experience exchanges in course of time in case these spins are directed oppositely.

The theory can be generalised from two to an arbitrary number N of electrons.⁴

* for this type of particles.

⁴ Cf. P.A.M. Dirac, Proc. Roy. Soc. Lond., Ser. A 123, 714 (1929); J.C. Slater, Phys. Rev. 34, 1293 (1929); for reviews see Rapport du Congrès de Solvay (1930), report by Pauli, in particular I, § 4; also M. Born, Z. Physik 64, 729 (1930); Ergebn. exakt. Naturw. 10, 387 (1931).

We shall not go here into the proofs but only sketch the results. It is advisable, then, not to start from the general theory of the representations of the permutation group, since even with respect to the symmetry of the eigenfunctions in the position coordinates alone, only a small part of all the possible representations can actually occur, as a result of the exclusion principle. It is advisable to substitute for the $u_n(q)$ in the determinant (14.8), expressions of the form (14.11) and (14.12) and to order the expressions thus obtained suitably. It turns out that as a result of the application of the operator $s^2 = \sum_{n=1}^N \sum_{k=1}^3 s_{nk}^2$ to the spin eigenfunctions, which is equivalent to multiplying the spin eigenfunctions by a number of the form $s(s+1)$, the symmetry character of the spin and hence, (on account of the exclusion principle), also the position eigenfunctions are uniquely determined. From this it follows that in the Russell-Saunders coupling ($H^{(2)}$ is small compared to $H^{(1)}$), even for N electrons the terms break up into different multiplets with energies differing by linear combinations of the "exchange integrals".

It may be remarked that if the earlier considerations regarding collisions are applied to the symmetry of eigenfunctions under interchange of the position coordinates alone we get the result that even on neglecting the spin forces in a collision of an electron with an atom, the latter can go over into a state with a multiplicity different from that of the initial state. For, here other symmetry classes can appear in addition to the symmetric and antisymmetric classes.

We have still to consider the statistical applications that are characteristic of systems of many identical particles belonging to a definite symmetry class (symmetric or antisymmetric). To this end, we consider a large number of free particles enclosed in a volume V so that the eigenfunctions are standing plane waves. The number of stationary states of a particle which lies in the volume element $(p_k, p_k + dp_k)$ ($k = 1, 2, 3$) of momentum space is then

$$Z = V \frac{1}{h^3} dp_1 dp_2 dp_3.$$

For particles with spin there is an additional weight factor $g = 2s + 1$ ($g = 2$ for electrons, protons, etc.) but we shall disregard it here for the sake of simplicity. It is then useful to consider not an individual state but a group of Z states of the kind considered, where Z is a large number. On the other hand, the energy of the particles within the group should vary so little that the Boltzmann factor $e^{-E/kT}$ is noticeably constant for the group of states. We can then ask for the a priori probability W that N particles lie in this group of states. This probability is given, according to general principles, by the number of stationary states of the total system of N particles, in which each particle has a momentum between p_k and $p_k + dp_k$. This number (of stationary states) depends on the symmetry class. For distinguishable particles this a priori probability is simply

$$W = Z^N. \quad (14.17)$$

For particles with symmetric states things are essentially different. To the two cases (i) the first particle is in state 1 and the second particle is in state 2 and (ii) the second particle is in state 1 and the first is in state 2, corresponds only one state of a system consisting of two particles. Therefore, it is a priori equally probable that of the two particles present one is in state 1, the other is in state 2 or both are in the same state, either 1 or 2. For distinguishable particles, on the contrary, the probability for each of the latter possibilities would have been only half as large as the probability of the

first. In general, for N particles, every symmetric eigenfunction is uniquely characterised by the number of particles found in each of the Z states of the group. We shall call such a statement, a "complexion". We find the number of such complexions to be equal to

$$W = \frac{(N + Z - 1)!}{N!(Z - 1)!} \text{ (symmetric states).} \quad (14.18a)$$

In the case of antisymmetric eigenfunctions (exclusion principle) we must drop from these complexions all those in which more than one particle are in the same state. We then find

$$W = \frac{Z!}{N!(Z - N)!} \text{ (antisymmetric states),} \quad (14.18b)$$

where, necessarily $N \leq Z$. If we have many groups of states with single particles, the number of the corresponding states of the total system is equal to the product of the numbers W for the individual groups. [For distinguishable particles there is an additional factor $N!/(N_1!N_2!\dots)$ where the N_n denote the numbers of particles in the individual groups while $N = \sum_n N_n$ is the total number of particles. For particles with

spin, this W is to be simply raised to the power g , if N is the number of particles in a definite spin state.]

Before the symmetry properties of the eigenfunctions of N identical particles were known, the enumeration method for symmetric states appeared to be a particularly hypothetical prescription. It was introduced by Bose;⁵ the interpretation of radiation as a gas consisting of corpuscular light-quanta led to correct results. It was extended to gases consisting of material particles by Einstein.⁵ We, therefore, speak of the *Bose-Einstein statistics*. But we are not dealing with a new kind of statistics* since we now know that the a priori probabilities are always proportional to the number of the stationary states of the total system concerned. We shall, therefore, rather speak of the statistics of symmetric states. It is to be used for those material particles which possess such symmetric states, as e.g., the α -particles. For particles which obey the exclusion principle, the corresponding statistical results were given by Fermi.⁵ Dirac⁵ also arrived at the same statistics basing his ideas on the antisymmetric eigenfunctions. We, therefore, speak of *Fermi-Dirac statistics*; but we shall prefer to call it the statistics of antisymmetric states.⁶

We shall now discuss an application of these two types of statistics, since it can be formulated without going into problems in heat. Here we deal with the fluctuations in the particle number and the energy in a sub-volume for an ensemble of N free parti-

⁵ See footnote I above.

* This and the following statements need not be taken to mean any belittling of the novelty and importance of the two statistics, viz., Bose-Einstein and Fermi-Dirac. The relation between the spin-statistics and type of wave function is only an observation and it required the full apparatus of relativistic quantum field theory and the complex Lorentz group to prove the spin-statistics theorem – see for example, R.F. Streater and A.S. Wightman, "PCT, Spin and Statistics, and All That", Benjamin, New York (1964); W. Pauli in "Niels Bohr and the Development of Physics", Pergamon Press, London (1955).

⁶ On the further thermodynamic consequences and applications of these statements cf. the monograph of L. Brillouin, *Die Quantenstatistik*, Berlin (1931); also P. Jordan, *Statistische Mechanik auf quantentheoretischer Grundlage*, Braunschweig (1933).

cles lying in a definite velocity interval. The relations concerning the particle number are simpler and will be treated first. If x_r is the abbreviation for the three space co-ordinates of the n^{th} particle, then

$$n(x_1, \dots, x_N) = \sum_{r=1}^N \int d^3x \delta(x - x_r) \quad (14.19)$$

is equal to the number of those co-ordinates x_r which lie in the partial volume v considered and over which integrations are to be performed. Then

$$\bar{n} = \int n(x_1, \dots, x_N) |U(x_1, \dots, x_N)|^2 d^3x_1 \dots d^3x_N,$$

$$\bar{n}^2 = \int n^2(x_1, \dots, x_N) |U(x_1, \dots, x_N)|^2 d^3x_1 \dots d^3x_N,$$

where $U(x_1, \dots, x_p)$ denotes the eigenfunction. The above integrations can be carried out. If we are dealing with a group of Z states where Z is a large number and if each state of the ensemble is considered to have equal probability and if further we confine ourselves, for the sake of simplicity, to the case that the volume v is small compared

to the total volume, then we have, with $z = Z \frac{v}{V}$, the well-known result:⁷

$$\overline{(\Delta n)^2} = \bar{n}^2 - (\bar{n})^2 = \bar{n} + \frac{\bar{n}^2}{z} \quad \text{for symmetric states,} \quad (14.20 \text{a})$$

$$\overline{(\Delta n)^2} = \bar{n}^2 - (\bar{n})^2 = \bar{n} - \frac{\bar{n}^2}{z} \quad \text{for antisymmetric states.} \quad (14.20 \text{b})$$

For independent particles, on the contrary $\overline{(\Delta n)^2} = \bar{n}$, as is well known. It is important to notice here that quantities of the order of magnitude $1/z$ relative to the terms written are neglected.

In considering the corresponding problem of the energy in a partial volume we have to be more careful since the knowledge of the momentum of a particle carries with it an uncertainty regarding the position of the particle. The measurement of the particle numbers in a sub-volume follows simply by determining the positions of all particles whose numerical values lie within the required limits. In contrast, the energy can be determined only by introducing a wall (or potential barrier) or some similar external influence which determines the boundaries of the sub-volume. The energy in the sub-volume after this influence is then essentially the same as the energy that it contained in a volume (before the influence), whose boundaries are indeterminate by an order of magnitude of the wave-length of the matter waves. Only when the sub-volume is large compared to this mean wave-length does the question regarding the energy in the sub-volume have a specific meaning. According to Heisenberg⁸ this must be taken into account if certain singularities which occur when we write the energy in the sub-volume, say, in the form

$$E = \frac{1}{2m} \sum_r \mathbf{p}_r D(\mathbf{x}_r) \mathbf{p}_r, \quad D(\mathbf{x}) = \int_v \delta(\mathbf{x} - \mathbf{x}') d^3x'$$

are to be avoided. The singularity vanishes, however, if we replace the function $D(x)$ by a continuous function, i.e. if we introduce a weight function which, outside the region v considered, falls off steeply but continuously:

$$\int G(x') \delta(x - x') d^3x' = G(x).$$

⁷ I thank Dr. R. Peierls for carrying out the calculation according to this method.

⁸ W. Heisenberg: Leipzig, Akad., math.-phys. Kl. 83, 3 (1931).

We also construct the expression

$$E = \frac{1}{2m} \sum_r \mathbf{p}_r G(\mathbf{x}_r) \mathbf{p}_r .$$

Then we have, analogous to (14.20):

$$(\Delta E)^2 = \overline{E^2} - \overline{E}^2 = \frac{1}{2m} \overline{\mathbf{p}^2} \overline{E} + \frac{\overline{E^2}}{z} \quad \text{for symmetric states,} \quad (14.21a)$$

$$(\Delta E)^2 = \overline{E^2} - \overline{E}^2 = \frac{1}{2m} \overline{\mathbf{p}^2} \overline{E} - \frac{\overline{E^2}}{z} \quad \text{for antisymmetric states.} \quad (14.21b)$$

We wish to mention here a peculiar mathematical method due to Jordan and Klein⁹ for the case of symmetric states and due to Jordan and Wigner¹⁰ for the antisymmetric states. This is called the method of second quantisation of waves in ordinary three-dimensional space. This method originated from a consideration of the analogy between material particles with symmetric states on the one hand and the quanta of radiation (photons) on the other. It is doubtful whether this method represents an actual deep lying physical analogy and it can also be shown that all the results of quantum mechanics can be obtained without using this method. However, as a method of calculation at least, we would like to give it here.

In section 9 we have denoted by $c_n^* c_m$ the probability that a particle is found in a state described by the eigenfunction $u_n(x)$ (see eq. (9.1) and ff). We now introduce (time-dependent) operators a_n^* and a_m which satisfy the C.R.

$$a_n a_m^* - a_m^* a_n = \begin{cases} 0 & \text{for } n \neq m, \\ 1 & \text{for } n = m. \end{cases} \quad (14.22a)$$

while

$$a_n a_m - a_m a_n = 0; \quad a_n^* a_m^* - a_m^* a_n^* = 0. \quad (14.22a')$$

Here a^* denotes the operator which is the Hermitian conjugate of a . Then it is easy to see that the eigenvalues of

$$a_m^* a_m = N_m \quad (14.23)$$

are the integers 0, 1, 2, ... If we write N_m as a diagonal matrix, then the matrices of a_m^* and a_m are

$$(a_m^*)_{N_m N_m'} = \begin{cases} \sqrt{N_m} & \text{for } N_m' = N_m - 1, \\ 0 & \text{otherwise} \end{cases} \quad (14.24a)$$

$$(a_m)_{N_m N_m'} = \begin{cases} \sqrt{N_m + 1} & \text{for } N_m' = N_m + 1, \\ 0 & \text{otherwise.} \end{cases} \quad (14.24a')$$

Thus a_m^* as an operator applied to a function $f(N_m)$ takes it to $\sqrt{N_m + 1} f(N_m + 1)$; similarly $a_m f(N_m)$ goes over to $\sqrt{N_m} f(N_m - 1)$. If we set

$$a_m^* = \sqrt{N_m} A_m^*, \quad a_m = A_m / \sqrt{N_m}, \quad (14.25)$$

where

$$A_m^* A_m = 1,^{11} \quad (14.26)$$

then

$$\begin{aligned} A f(N_m) &= f(N_m - 1), \\ A^* f(N_m) &= f(N_m + 1). \end{aligned} \quad (14.27)$$

Similar C.R. can be set up, according to Jordan and Wigner, for particles with antisymmetric states, where the N_n can have only the values 0 or 1. Following these authors, we can set

$$a_n a_m^* + a_m^* a_n = \begin{cases} 0 & \text{for } n \neq m, \\ 1 & \text{for } n = m, \end{cases} \quad (14.22b)$$

$$a_n a_m + a_m a_n = 0, \quad a_n^* a_m^* + a_m^* a_n^* = 0 \quad \text{for } m \neq n, \quad (14.22b')$$

⁹ P. Jordan and O. Klein, Z. Physik 45, 751 (1927).

¹⁰ P. Jordan and E. Wigner, Z. Physik 47, 631 (1928).

¹¹ It is often written $\Delta_m = e^{i\Theta_m \hat{n}}$ in order to satisfy eq. (14.26) identically.

where, again we have

$$a_n^* a_n = N_n. \quad (14.23)$$

Further, the matrices will now be

$$(a_n^*)_{1,0} = \epsilon_n, \quad (a_n)_{0,1} = \epsilon_n, \quad (a_n^*)_{N_n N_n'} = (a_n)_{N_n' N_n}, \quad (14.24\text{b})$$

where $\epsilon_n = \pm 1$ with the sign still to be determined and dependent on n . To determine the sign, the states n must be considered to be ordered in a particular way. Then we can set

$$\epsilon_n = \prod_{m \leq n} (1 - 2N_m), \quad (14.28)$$

which is equal to +1 or -1 according as the number of occupied states, which lie before the particular state considered, is even or odd. Then we have to set

$$\left. \begin{aligned} a_n^* f(N_1 \dots 0_n \dots) &= \epsilon_n (N_1 \dots 0_n \dots) f(N_1 \dots 1_n \dots) \\ &= -\epsilon_n (N_1 \dots 1_n \dots) f(N_1 \dots 0_n \dots), \\ a_n^* f(N_1 \dots 1_n \dots) &= 0. \end{aligned} \right\} \quad (14.29)$$

$$\left. \begin{aligned} a_n f(N_1 \dots 0_n \dots) &= 0, \\ a_n f(N_1 \dots 1_n \dots) &= \epsilon (N_1 \dots 0_n \dots) f(N_1 \dots 0_n \dots) \\ &= -\epsilon_n (N_1 \dots 1_n \dots) f(N_1 \dots 0_n \dots). \end{aligned} \right\} \quad (14.29')$$

We can go over easily from the operators a_n and a_n^* to the ψ -operators themselves according to

$$\Psi(q) = \sum_n a_n(t) u_n(q); \quad \Psi^*(q) = \sum_n a_n^*(t) u_n^*(q). \quad (14.30)$$

where q combines the position and spin co-ordinates and u_n and u_n^* are c-numbers and form an orthonormal system. The latter property has the consequence that analogous to (14.22a, b) the ψ 's satisfy the C.R.:

$$\Psi(q) \Psi^*(q') \mp \Psi^*(q') \Psi(q) = \delta(q - q') \mathbf{1}, \quad (14.31)$$

$$\Psi^*(q) \Psi^*(q') \mp \Psi^*(q') \Psi^*(q) = 0, \quad \Psi(q) \Psi(q') \mp \Psi(q') \Psi(q) = 0. \quad (14.31')$$

Here the upper or the lower sign holds according as one is dealing with a symmetric or an antisymmetric state respectively. $\delta(q - q')$ stands for $\delta(x - x')\delta_{\mu\mu'}$ where $\delta_{\mu\mu'}$ is the Kronecker delta symbol for the discrete spin co-ordinates.

As an application of this method, we shall calculate again the density and energy fluctuations, for which we have to construct the operators

$$n = \sum_{s_1 s_2} \int \psi \psi^* d^3 x,$$

$$E = \sum_{s_1 s_2} \int \frac{\hbar^2}{2m} \frac{\partial \psi}{\partial x_s} \frac{\partial \psi^*}{\partial x_s} d^3 x$$

and the mean values (expectation values) of n , n^2 and E , E^2 with respect to the group of states considered. The result is the same as the calculation in configuration space.¹²

This is a special case of the general equivalence of the method of quantised eigen-vibrations and the method of configuration space, which – as the authors mentioned have shown – can be extended even to the problem of identical, interacting particles. The Hamiltonian operator is

$$H = \frac{1}{2m} \sum_r \left[-\hbar^2 \frac{\partial^2}{\partial x_r^2} + \sum_r V_r(q_r) + \sum_{r < s} \Omega(q_r, q_s) \right]. \quad (14.32)$$

Here the external forces are represented by $V(q_r)$ while the function $\Omega(q_r, q_s)$ which depends on a pair of particles, describes the interaction between the particles. In the case of Coulomb electrostatic forces

¹² In this method the energy density of free point masses will be formally analogous to the energy density of a vibrating string with quantised eigen-vibrations. This latter system was first investigated as regards its fluctuation properties by M. Born, W. Heisenberg and P. Jordan, Z. Physik 35, 557 (1925).

$H_{rs} = \frac{e^2}{r_{rs}}$; with a view to later generalisations which will include the magnetic interactions of particles, we shall allow V and Ω to depend on the spin co-ordinates also. If we take $\rho(q) = \psi^*(q)\psi(q)$ to be the classical charge cloud, we can write the interaction energy of the particles r and s classically as

$$\iint dq_r dq_s \Omega(q_r, q_s) \rho(q_r) \rho(q_s).$$

Then the Hamiltonian is defined by

$$H = \frac{1}{2m} \sum_{s_r} \int \left[\hbar^2 \frac{\partial \psi^*}{\partial x_r} \frac{\partial \psi}{\partial x_r} + V(q_r) \psi^* \psi \right] dx_r + \left. \quad \quad \quad \right\} \\ + \sum_{s_r, s_s} \iint \psi^*(q') \psi^*(q) \Omega(q, q') \psi(q') \psi(q) dq_r dq_s, \quad (14.32')$$

where ψ^* and ψ are the operators employed above which can be expressed according to eqs. (14.30) by the operators a_r^*, a_s^* arising in the expansion of ψ in terms of any orthogonal system. If we introduce the numbers N_n of the particles in the states n defined by this system as variables of a wave function $\Phi(N_1, N_2, \dots, t)$, on which the operator defined by (14.32') acts, then we can set up a wave equation of the type

$$-\frac{\hbar}{i} \frac{\partial \Phi}{\partial t} = H \Phi(N_1, N_2, \dots, t). \quad (14.33)$$

It turns out that the results arising from this wave equation coincide completely with the results arising from the wave equation in configuration space with the Hamiltonian (14.32). This holds for particles with symmetric as well as antisymmetric states.¹³ For this agreement the sequence of factors in eq. (14.32') is essential.

If many different kinds of particles are present (e.g., electrons and protons), we can introduce for each kind of particle a particular ψ -operator. These operators commute if they belong to different types of particles.

This, in brief, is the method of quantised eigen-vibrations. It is to be stressed that in spite of the formal mathematical analogy, an essential physical difference exists between the transition from the classical quantities p and q to the corresponding quantum mechanical operators, on the one hand, and the transition from the functions in ordinary space ψ^* and ψ to the operators ψ^* and ψ , on the other. For, even the wave functions ψ^* and ψ are symbolic quantities, which are themselves not observable but contain the quantum of action.

¹³ Cf. for the proof, besides the papers already cited, also V. Fock, Z. Physik 75, 622 (1932) and the book of W. Heisenberg, loc. cit.

CHAPTER VIII

Semiclassical Theory of Radiation

15. Treatment of the Radiation Processes Based on the Correspondence Principle

Historically, the process of light emission played, as is well known, an important role in laying the foundation of Heisenberg's Matrix Mechanics in which the matrix elements of the electric moment of the atom were associated directly with the electric field strengths of the light emitted in the corresponding processes. This association was based on classical electrodynamics. The formalism was then extended to dispersion phenomena by Born, Heisenberg, and Jordan.¹ The corresponding treatment using the Schrödinger equation was given by Klein.² But it turned out that special prescriptions which apparently did not follow from the general principles of quantum mechanics had to be introduced for reaching conclusions regarding the above-mentioned relationship between the electric moment of the atom and the outgoing radiation. This deficiency was removed only in the subsequent quantum mechanical treatment of light waves by Dirac.* Since the resulting theory of Dirac leads to particular difficulties connected with the problem of the structure of the electron itself, the original treatment is also of particular interest. This method leans more closely on the corresponding classical theory as the electromagnetic field is not quantised. We shall formulate the theory in such a way that the transition to Dirac's theory of radiation can be effected in the most direct manner possible. Here, however, no restrictive assumptions will be made regarding the number of electrons in the atom and the ratio of their wave-lengths to atomic dimensions.

Let us consider *classically* a system of particles about which specific statistical data are available, viz., to each configuration of the positions $x_k^{(a)}$ ($k = 1, 2, 3; a = 1, \dots, N$) of the particles lying in the interval $(x_k^{(a)}, x_k^{(a)} + dx_k^{(a)})$, a probability $p(x_1^{(1)}, \dots, x_N^{(N)}; t)$ for this configuration and an associated mean current $i_k^{(a)}(x_1^{(1)}, \dots, x_N^{(N)}; t)$ of the particles are given. (The symbol a labels the particles.) Then

$$\begin{aligned}\bar{\rho}^{(a)} &= \int \rho d^3x^1 \dots d^3x^{(a-1)} d^3x^{(a+1)} \dots d^3x^{(N)}, \\ \bar{i}_k^{(a)} &= \int i_k^{(a)} d^3x^1 \dots d^3x^{(a-1)} d^3x^{(a+1)} \dots d^3x^{(N)}\end{aligned}\quad (15.1)$$

are the values of the density and current of particle (a) at the space-point $x_k^{(a)}$ at

¹M. Born, W. Heisenberg and P. Jordan, Z. Physik 35, 557 (1926).

²O. Klein, Z. Physik 41, 407 (1927).

*See P.A.M. Dirac, The Principles of Quantum Mechanics, Oxford, Fourth Edition (1959), Ch. 10.

time t averaged over the positions of the other particles. The average value of the scalar potential Φ_0 and the vector potential Φ_k at the point P with the co-ordinates x_p at time t are, according to classical electrodynamics, given by

$$\left. \begin{aligned} \Phi_0(x_p; t) &= \sum_{a=1}^N \int \frac{\bar{\rho}^{(a)}(x_Q; t - \frac{r_{PQ}}{c})}{r_{PQ}} d^3 x_Q^{(a)}, \\ \Phi_k(x_p; t) &= \sum_{a=1}^N \int \frac{\bar{i}_k^{(a)}(x; t - \frac{r_{PQ}}{c})}{r_{PQ}} d^3 x_Q^{(a)}. \end{aligned} \right\} \quad (15.2)$$

These expressions simplify, if we consider the distance of the point P from the sources Q to be large compared to the dimensions of the region within which $\bar{\rho}^{(a)}$ and $\bar{i}_k^{(a)}$ are significantly different from zero. We shall confine ourselves in this section to the wave-zone of P . Introducing the distance R_p of the external point from a fixed point O in the system, we can set, in a well-known way,

$$r_{PQ} = R_p - (\vec{x}_Q \cdot \vec{n}), \quad (15.3)$$

where \vec{n} denotes a unit vector in the direction from O to P and \vec{x}_Q , the vector along OQ . We shall concentrate on the terms proportional to $1/R_p$ in the wave-zone, both in the expression for the potentials and in the field strengths following from them. From eqs. (15.2) and (15.3) then follows that

$$\left. \begin{aligned} \Phi_0(x_p; t) &= \frac{1}{R_p} \sum_{a=1}^N \int \bar{\rho}^{(a)}(x_Q; t - \frac{R_p}{c} + \frac{1}{c}(\vec{x}_Q \cdot \vec{n})) d^3 x_Q^{(a)}, \\ \Phi_k(x_p; t) &= \frac{1}{R_p} \sum_{a=1}^N \int \frac{1}{c} \bar{i}_k^{(a)}(x_Q; t - \frac{R_p}{c} + \frac{1}{c}(\vec{x}_Q \cdot \vec{n})) d^3 x_Q^{(a)}. \end{aligned} \right\} \quad (15.4)$$

In changing over to the field strengths we have to remember that in the differentiations with respect to $(x_p)_k$, R_p is to be considered a constant in the approximation considered here. From

$$\begin{aligned} \frac{\partial}{\partial x_{k,P}} \int f(x_Q; t - \frac{r_{PQ}}{c}) d^3 x_Q &= -\frac{1}{c} \frac{\partial}{\partial t} \int f(x_Q; t - \frac{r_{PQ}}{c}) \frac{\partial r_{PQ}}{\partial x_{k,P}} d^3 x_Q \\ &= +\frac{1}{c} \frac{\partial}{\partial t} \int f(x_Q; t - \frac{r_{PQ}}{c}) \frac{\partial r_{PQ}}{\partial x_{k,Q}} d^3 x_Q \end{aligned}$$

and eq. (15.3), we have in the wave-zone

$$\frac{\partial}{\partial x_{k,P}} \int f(x_Q; t - \frac{r_{PQ}}{c}) d^3 x_Q = -\frac{1}{c} n_k \frac{\partial}{\partial t} \int f(x_Q; t - \frac{r_{PQ}}{c}) d^3 x_Q.$$

We then obtain for the parts of the field strengths proportional to $1/R_p$

$$\left. \begin{aligned} \vec{E} &= -\frac{1}{c} \frac{\partial \vec{\Phi}}{\partial t} - \text{grad } \Phi_0 = -\frac{1}{c} \frac{\partial \vec{\Phi}}{\partial t} + \vec{n} \frac{1}{c} \frac{\partial \Phi_0}{\partial t}, \\ \vec{H} &= \text{rot } \vec{\Phi} = -\left[\vec{n}, \frac{1}{c} \frac{\partial \vec{\Phi}}{\partial t} \right]. \end{aligned} \right\} \quad (15.5)$$

While \vec{H} is perpendicular to \vec{n} , \vec{E} seems to contain a longitudinal part, i.e. a part parallel to \vec{n} . On the basis of the continuity equation for $\bar{i}^{(a)}$ and $\bar{\rho}^{(a)}$ follows easily the

relation³

$$\frac{1}{c} \frac{\partial \Phi_0}{\partial t} = \frac{1}{c} (\vec{n} \frac{\partial \vec{\Phi}}{\partial t}), \quad (15.6)$$

which is valid in the wave-zone and which has, on account of (15.5), the consequence that the longitudinal part of the field strength vanishes in the wave-zone

$$(\vec{\mathcal{E}} \cdot \vec{n}) = 0 \quad (15.6')$$

[i.e. $(\vec{\mathcal{E}} \cdot \vec{n})$ vanishes faster than $1/R_p$]. On introducing the transversal component

$$\vec{\Phi}_{tr} = \vec{\Phi} - \vec{n} (\vec{n} \cdot \vec{\Phi}) = \frac{1}{R_p} \sum_{n=1}^N \frac{1}{c} \int \vec{i}_{tr}(x_Q; t - \frac{R_p}{c} + \frac{1}{c} (\vec{x}_Q, \vec{n})) d^3 x_Q^{(n)} \quad (15.7)$$

of the vector potential, we can, therefore, write the relations (15.5), according to (15.6), as

$$\vec{\mathcal{E}} = - \frac{1}{c} \frac{\partial \vec{\Phi}_{tr}}{\partial t}; \quad \vec{\mathcal{H}} = - \left[\vec{n}, \frac{1}{c} \frac{\partial \vec{\Phi}_{tr}}{\partial t} \right] = [\vec{n}, \vec{\mathcal{E}}]. \quad (15.8)$$

The Poynting vector will be

$$\vec{S} = \frac{c}{4\pi} [\vec{\mathcal{E}}, \vec{\mathcal{H}}] = \frac{c}{4\pi} \vec{n} \mathcal{E}^2 = \frac{c}{4\pi} \vec{n} \mathcal{H}^2. \quad (15.9)$$

We have to face now the problem of carrying over these results of classical theory to quantum mechanics. In quantum mechanics each state of the system is in principle described statistically by any solution

$$\psi(x_1 \dots x_{3N}; t) = \sum_n c_n u_n(x_1 \dots x_{3N}; t)$$

of the corresponding wave equation when the u_n is any orthonormal system of particular solutions of the wave equation. At first it might be thought that this expression for ψ is to be substituted in the expression for the current \vec{i} which is bilinear in ψ^* and ψ , and then $\vec{\Phi}_{tr}$ and $\vec{\mathcal{E}}, \vec{\mathcal{H}}$ have to be formed according to eqs. (15.7) and (15.8). This would yield the mean value (expectation value) of the potential and of the field strengths at a point. *The measurement of the radiation emitted by the system, however, never consists in a determination of the expectation values of the field strengths.* These expectation values vanish, e.g., for a stationary state for which \vec{i} is independent of time. *Rather, one is always interested in determining the expectation values of expressions quadratic in the field strengths.* Later, we shall see that in processes involving low light intensity in which only a small and well-defined number of photons play a role, the field strengths themselves are to be considered unmeasurable (apart from the trivial stipulation that the time average of their expectation values vanishes). Of course, it is not only possible to measure the time average of the square of the total field strengths at a space-point but also the time average of the square of the amplitudes of any Fourier components of $\vec{\mathcal{E}}$ or $\vec{\mathcal{H}}$. This is possible because the photographic plates, ionisation chambers, absorbing atoms, etc., by which the light is detected can react with different strengths for different

³ This is connected with the fact that due to the continuity equation, the expression (15.2) satisfies the condition

$$\frac{1}{c} \frac{\partial \Phi_0}{\partial t} + \operatorname{div} \vec{\Phi} = 0$$

which goes into (15.6) in the wave-zone.

frequencies. Here a *temporal* Fourier-decomposition of $\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ is made according to

$$\vec{\mathcal{E}}(x_p; t) = \sum_{\omega} \vec{\mathcal{E}}(\omega; x_p) e^{i\omega t}; \quad \vec{\mathcal{H}}(x_p; t) = \sum_{\omega} \vec{\mathcal{H}}(x_p; \omega) e^{i\omega t},$$

where the sum can also be replaced by an integral under some circumstances and

$$\vec{\mathcal{E}}(-\omega) = \vec{\mathcal{E}}^*(\omega); \quad \vec{\mathcal{H}}(-\omega) = \vec{\mathcal{H}}^*(\omega);$$

i.e. changing ω to $-\omega$ amounts to taking the complex conjugates of the amplitudes. At present we need not investigate the question as to how the spatial dependence of the expectation values of $\vec{\mathcal{E}}_o^2$ can be determined by measurement. At any rate, the measurement can be done under certain circumstances in spatial regions which are small compared to the wave-length of the light, as one knows, e.g., from the well-known experiments on standing light waves.

Since \vec{t} is bilinear in ψ and ψ^* the expectation value of every quantity F linear in the components of the field strengths can be represented in the form

$$F(x_p; t) = \sum_{n,m} c_n^* F_{n,m}(x_p; t) c_m, \quad (15.10)$$

if $F_{n,m}$ are matrix elements which arise by substituting $\psi^* = v_n^*$ and $\psi = v_n$ in \vec{t} . Then the expectation value of F^2 is

$$(F^2) = \sum_{n,m} c_n^* (F^2)_{n,m}(x_p; t) c_m = \sum_{n,m} c_n^* \sum_l F_{n,l}(x_p; t) F_{l,m}(x_p; t) c_m,$$

as was shown in Sec. 9. Further the time average of the expectation value (F^2) will be

$$\begin{aligned} (\bar{F}^2) &= \sum_n c_n^* \sum_l \sum_{\omega} F_{n,l}(\omega; x_p) F_{l,m}(-\omega; x_p) c_m \\ &= \sum_n c_n^* \sum_l \sum_{\omega>0} [F_{n,l}(\omega; x_p) F_{l,m}(-\omega; x_p) + F_{n,l}(-\omega; x_p) F_{l,m}(\omega; x_p)] c_m. \end{aligned} \quad (15.11)$$

At this point a certain ambiguity enters in this interpretation based on the correspondence principle, since the $F_{n,m}(\omega; x_p)$ need not be Hermitian but in general we have only the relation

$$F_{n,m}^*(\omega) = F_{m,n}(-\omega). \quad (15.12)$$

This ambiguity was removed by a special prescription formulated by Klein the meaning of which is not clear in this approach but which is necessary in order to have agreement with experience and indeed even with the validity of the energy conservation law in individual emission or scattering processes. For the wave-zone, i.e. outside the emitting or scattering system itself, this prescription runs as follows: *Prescription I.* Each quantity F considered is divided into $F^{(+)}$ and $F^{(-)}$ according to⁴

$$F = F^{(+)} + F^{(-)}, \quad (15.13)$$

$$F^{(+)} = \sum_{\omega>0} F(\omega; x_p) e^{i\omega t}; \quad F^{(-)} = \sum_{\omega<0} F(\omega; x_p) e^{+i\omega t} = \sum_{\omega>0} F(-\omega; x_p) e^{-i\omega t}. \quad (15.14)$$

Therefore,

$$F_{n,m}^+ = \sum_{\omega>0} F_{n,m}(\omega; x_p) e^{i\omega t}; \quad F_{n,m}^- = \sum_{\omega<0} F_{n,m}(\omega; x_p) e^{i\omega t} = \sum_{\omega>0} F_{n,m}(-\omega; x_p) e^{-i\omega t}, \quad (15.14')$$

⁴ We have neglected here $F^{(0)} = \bar{F}$ as we are not interested in static fields.

and the time average of the expectation value of the classical quantity F^2 is replaced by $\bar{F}^+ F^-$:⁵

$$(\bar{F}^2) \rightarrow 2(\bar{F}^+ F^-) \quad (15.15)$$

and correspondingly

$$\begin{aligned} F(\omega) F(-\omega) + F(-\omega) F(\omega) &\rightarrow 2 F(\omega) F(-\omega) \\ = 2 \sum_{n,m} c_n^* \sum_l F_{n,l}(\omega) F_{l,m}(-\omega) c_m. \end{aligned} \quad \left. \right\} \quad (15.15')$$

This prescription is to be employed in the same way for emission and scattering. In the former case, the orthogonal solutions of the wave equation of the unperturbed system are to be substituted for v_n and in the latter case, the orthogonal solutions⁶ of the system perturbed by the external radiation. Actually which (time-dependent) orthogonal system is to be used remains completely arbitrary.

As an application of these general ideas we consider light emission more closely and choose for the v_n the solutions corresponding to the stationary states of the unperturbed system

$$u_n(x_1, \dots, x_{3N}) e^{-\frac{iE_n}{\hbar} t},$$

which depend exponentially on time. According to (15.7) the matrices of the transversal components of the vector potential of the emitted light will be

$$(\vec{\Phi}_{tr})_{n,m} = \frac{e^{i\tau_{n,m}t}}{R} \frac{(-e)}{c} \sum_{a=1}^N \int \vec{i}_{tr}^{(a)}(u_n^*, u_m) e^{i(\vec{k}_{n,m} \vec{x}^{(a)})} d^3x^{(a)}. \quad (15.16)$$

Here the emission frequency is

$$\nu_{n,m} = \frac{E_n - E_m}{\hbar}, \quad (15.17)$$

and the wave vector $\vec{k}_{n,m}$ of the emitted light is

$$\vec{k}_{n,m} = \frac{\nu_{n,m}}{c} .$$

The factor $(-e)$, being the charge of the electron, has been added in eq. (15.16) because, for solutions which are normalised to 1, the expression for \vec{i} given earlier (eq. (15.1)) denotes the particle current. According to (4.18) and (5.14) we have

$$\begin{aligned} \vec{i}_{tr}^{(a)}(u_n^*, u_m) &= \frac{\hbar}{2m} \int d^3x^{(1)} d^3x^{(2)} \dots d^3x^{(a-1)} d^3x^{(a+1)} \dots d^3x^{(N)} \times \\ &\times \frac{1}{i} \left(u_n^* \frac{\partial u_m}{\partial x_k^{(a)}} - u_m \frac{\partial u_n^*}{\partial x_k^{(a)}} \right), \end{aligned} \quad \left. \right\} \quad (15.18)$$

if we assume, for the sake of simplicity, that no static magnetic field is present. In the relativistic theory a different expression will be used for the current; but the relation (15.16) remains the same there also. By substituting (15.18) in (15.16), the Hermiticity of the matrix $(\Phi_{tr})_{n,m}$ follows. The transversality of the vector potential Φ is essential for this purpose.

⁵ For this cf. also G.C. Wick, Phys. Rev. **80**, 268 (1950).

⁶ Even for time-dependent Hamiltonians the orthogonality and normalisation of a system of solutions of the wave equation remain unaltered in time, according to Sec. 8, provided that the Hamiltonian is real.

According to (15.16) the decomposition of $\vec{\Phi}$ into $\vec{\Phi}^+$ and $\vec{\Phi}^-$ is very easy; namely

$$\left. \begin{array}{ll} (\vec{\Phi}_{\text{tr}})_{n,m}^{(+)} = (\vec{\Phi}_{\text{tr}})_{n,m} & \text{for } v_{n,m} > 0 (E_n > E_m), \\ (\vec{\Phi}_{\text{tr}})_{n,m}^{(+)} = 0 & \text{for } v_{n,m} < 0 (E_n < E_m), \\ (\vec{\Phi}_{\text{tr}})_{n,m}^{(-)} = 0 & \text{for } v_{n,m} > 0 (E_n > E_m), \\ (\vec{\Phi}_{\text{tr}})_{n,m}^{(-)} = (\vec{\Phi}_{\text{tr}})_{n,m} & \text{for } v_{n,m} < 0 (E_n < E_m). \end{array} \right\} \quad (15.19)$$

Using eqs. (15.9), (15.16), (15.18), (15.19) and the prescription (15.15) and introducing the abbreviation

$$\vec{C}_{n,m} = \frac{\hbar}{2m} i v_{n,m} \int d^3x^{(1)} \dots d^3x^{(N)} \sum_{a=1}^N e^{i \vec{k}_{n,m} \cdot \vec{x}^{(a)}} \frac{1}{i} \left(u_n^* \frac{\partial u_m}{\partial \vec{x}_{\text{tr}}^{(a)}} - u_m \frac{\partial u_n^*}{\partial \vec{x}_{\text{tr}}^{(a)}} \right), \quad (15.20)$$

we obtain for the energy radiated in the direction \vec{n} per solid angle $d\Omega$ and per unit time:

$$\vec{S} = \frac{c}{4\pi} \sum_m 2 [\vec{\Phi}^{(+)} \times \vec{\Phi}^{(-)}]_{m;m} = \vec{n} \frac{e^2}{c^3} \frac{1}{4\pi} 2 \sum_{m(E_m < E_n)} |\vec{C}_{n,m}|^2. \quad (15.21)$$

This is the energy emitted if initially only the state n was present. The constraint that we should sum over only those states m for which $E_m < E_n$, is due to the special prescription (15.15); if we had also taken $[\vec{\Phi}^{(-)} \times \vec{\Phi}^{(+)}]$, we would have found, in violation of the conservation law for energy, an emission which would have corresponded with transitions to states of energy higher than that of the initial state.

If initially we have a general wave-packet

$$\sum_n c_n u_n$$

and not just a single stationary state, then, according to (15.5), we have to form

$$|\vec{S}| = \frac{e^2}{c^3} \frac{1}{4\pi} \sum_{n,m} 2 c_n^* \left(\sum_l \vec{C}_{n,l} \vec{C}_{l,m} \right) e^{i v_{n,m} t} c_m \text{ where } E_l < E_n; E_l < E_m. \quad (15.22)$$

But on constructing the time average, all terms for which $v_{n,m} \neq 0$ vanish and hence E_n and E_m are different. In a degenerate system, of course, many states with the same energy $E_n = E_m$ can be present.

We mention, further, that from (15.20) a simple selection rule, which is valid rigorously for arbitrarily short wave-lengths (multipole radiation) follows. If both in the initial and final states the eigenfunctions are rotation-invariant, which according to Sec. 13 corresponds to vanishing angular momentum, then $C_{n,m}$ vanishes and along with it, the radiation also. For, if the co-ordinate system is rotated about the axis parallel to $\vec{k}_{n,m}$, then in this case the integrand retains its value as well as its form; on the other hand it transforms like a vector, on account of the differentiation with respect to \vec{x}_{tr} (changes, e.g., its sign on rotation through 90°) and both are simultaneously possible only if $\vec{C}_{n,m}$ vanishes. *Transitions between states with the angular momentum $J=0$ initially and $J=0$ finally under spontaneous emission of light are, therefore, strictly forbidden.* It is easily seen that this rule holds true even when spin (cf. Sec. 13) is taken into account, if, by J is understood the resultant of the orbital and spin angular momenta. This rule holds for the total orbital momentum L by itself only if its coupling to the spin angular momentum can be neglected.

Till now no assumption has been made about the ratio of the dimensions of the system to the wave-length of the emitted light. If this ratio is small then only for small values of $(\vec{k}_{n,m} \vec{x})$ are the eigenfunctions noticeably different from zero and the exponential function $e^{-i(\vec{k}_{n,m} \vec{x}^{(a)})}$ can be, with advantage, expanded in a power series. *The individual terms of this expansion correspond to dipole-, quadrupole-... radiation.* In particular we get the dipole radiation if $e^{-i(\vec{k}_{n,m} \vec{x}^{(a)})}$ is replaced by 1; this is equivalent to saying that in (15.7) the retardation term $(1/c)(\vec{x}_g \vec{n})$ in the time-argument of the current will be neglected. Since the matrix elements $\vec{x}_{n,m}$ of the co-ordinates are related to those of the current $\vec{I}_{n,m}$, on account of the continuity equation, through

$$i v_{n,m} \vec{x}_{n,m} = \vec{I}_{n,m}$$

(cf. eq. (4.16')), we can replace the dipole radiation eq. (15.20) by

$$(\vec{C}_{n,m})_{\text{Dipole}} = -v_{n,m}^3 \vec{x}_{tx,n,m}. \quad (15.23)$$

Therefore, according to (15.21), we have

$$|\vec{S}|_n, \text{Dipole} = \frac{e^2}{c^2} \frac{1}{4\pi} 2 \sum_{m(E_m < E_n)} v_{n,m}^4 |\vec{x}_{tx,n,m}|^2. \quad (15.24)$$

This relation served Heisenberg originally to define matrices in his matrix mechanics.

Dispersion can also be treated in a similar fashion. The only point is that in this case a preliminary perturbation calculation is necessary to determine the influence of the external field on the eigenfunction of the atom. We can make the calculation as if we were dealing with a classical, electromagnetic field varying with time in a specified way, and characterised by its vector potential $\Phi_k(x, y, z; t)$. Of course, the field of the source of the incident light need not be measurable classically, but the result as well as the corresponding quantisation of the radiation field justify this mode of treatment. In the case of plane waves

$$\Phi_k = \varphi_k^+ e^{i(vt - \vec{k} \cdot \vec{x})} + \varphi_k^- e^{-i(vt - \vec{k} \cdot \vec{x})}, \quad (15.25)$$

where

$$\varphi_k^- = (\varphi_k^+)^*, \quad (15.26)$$

i.e. φ_k^- is the complex conjugate of φ_k^+ . The time average of the square of the field strength is

$$\overline{\mathcal{E}^2} = v^2 2 \varphi_k^+ \varphi_k^- = 2v^2 |\varphi_k|^2. \quad (15.27)$$

For light waves it is always permissible to set the scalar potential equal to zero and to normalise the vector potential, according to

$$\sum_{k=1}^3 \frac{\partial \Phi_k}{\partial x_k} = 0 \quad (15.25')$$

i.e. assume it to be transversal. Then, according to (5.11), the perturbation⁷ function reads

$$\Omega = \frac{1}{2m} \frac{\hbar}{i} \sum_{a=1}^N \left\{ \frac{e}{c} 2 \sum_{k=1}^3 \Phi_k(x^{(a)}) \frac{\partial}{\partial x_k^{(a)}} + \frac{1}{2m} \frac{e^2}{c^2} \sum_{k=1}^3 \Phi_k^2(x^{(a)}) \right\}$$

⁷ We replace the charge $e^{(a)}$ occurring in (5.11) and e_k in (5.14) by the electron charge ($-e$).

If

$$\vec{i}_{a,k}^{(0)} = \frac{\hbar}{2m} \frac{1}{i} \left(\psi^* \frac{\partial \psi}{\partial x_k^{(a)}} - \psi \frac{\partial \psi^*}{\partial x_k^{(a)}} \right)$$

is the unperturbed current, then the part of the perturbation function which is linear in Φ_k , written in the matrix form, is given by

$$\Omega_{n,m}^{(1)} = \frac{e}{c} \left\{ \sum_{a=1}^N \sum_{k=1}^3 \Phi_k(x^{(a)}) \vec{i}_k^{(a)} \right\}_{n,m} \quad (15.28)$$

Further, according to (5.14) there is an additional perturbation term in \vec{T} proportional to Φ_k .⁸

$$\vec{i}_k^{(1)} = \frac{e}{mc} \vec{\Phi}(x^{(a)}) \psi^* \psi. \quad (15.29)$$

As a consequence of (15.7), both the additional terms, – the one in the Hamiltonian and the other in the current – give rise to matrix elements of the vector potential of the emitted light that are proportional to the amplitude of the incident light. We shall not discuss terms of higher order⁹ here. [It may be remarked that the form (15.28) of the perturbation function remains the same even in the relativistic theory, though the current operator there is different; in contrast, eq. (15.29) is not satisfied there.]

The general form of the matrix elements of the scattered radiation, which follows by using perturbation theory and the general formula (15.7) is the following:

$$(\vec{\Phi}_u)'_{n,m} = \sum_{k=1}^3 \{ \varphi_k^+ \vec{a}_{k;n,m} e^{i(v_{n,m} + v)} + \varphi_k^- \vec{b}_{k;n,m} e^{i(v_{n,m} - v)} \}, \quad (15.30)$$

if only expressions linear in φ_k of the incident wave are retained. The prime in (15.30) distinguishes the scattered radiation from the incident radiation. Further, the field strengths are obtained by differentiation with respect to time and division by c . On account of the Hermiticity of the Hamiltonian, the matrix $(\vec{\Phi}_{tr})'_{n,m}$ is itself Hermitian and hence we have [cf. (15.26)]

$$\vec{b}_{k;n,m} = \vec{a}_{k;n,m}^*. \quad (15.31)$$

In words, \vec{a}_k is not Hermitian but \vec{b}_k is the matrix which is the Hermitian conjugate of \vec{a}_k .

We can now formulate the general prescription (15.15) for constructing the expression for the radiated energy. If the initial state is n ,

$$S_n = \frac{c}{4\pi} \nu'^2 2 |\varphi_k|^2 \vec{a}_{k;n,m} \vec{b}_{k;n,m} = \frac{c}{4\pi} \nu'^2 2 |\varphi_k|^2 |\vec{a}_{k;n,m}|^2 \quad \left. \begin{array}{l} \\ \text{for } \nu' = \nu_{n,m} + \nu > 0. \end{array} \right\} \quad (15.32a)$$

or

$$S_n = \frac{c}{4\pi} \nu'^2 2 |\varphi_k|^2 \vec{b}_{k;m,n} \vec{a}_{k;m,n} = \frac{c}{4\pi} \nu'^2 2 |\varphi_k|^2 |\vec{a}_{k;m,n}|^2 \quad \left. \begin{array}{l} \\ \text{for } \nu' = \nu_{n,m} - \nu > 0. \end{array} \right\} \quad (15.32b)$$

If we had not used the special prescription for separating the quantities into terms

⁸ See the previous footnote.

⁹ Regarding the details of calculation cf. besides the paper of Klein referred to, also I. Waller, Naturwiss. 15, 969 (1927); Phil. Mag. 4, 1228 (1927) especially for the case of short wave-lengths.

corresponding to $e^{i\omega t}$ and $e^{-i\omega t}$ ($\omega > 0$), then the states m and n would not have been distinguished and we would have arrived at the expression $\frac{1}{2}(S_n + S_m)$ for the emergent radiation, for both the states. In the particular case $n = m$, $\nu = \nu'$ the prescription in question yields, however, nothing new so that scattering without change of frequency can be understood even without its use.

The general form of the expressions for $\vec{a}_{k,n,m}$ will not be discussed here.* Because of its importance we can refer here only to a special case in which the frequency ν of the incoming radiation is high compared to the work done in separating an electron from the system. It turns out that the terms in (15.30) give the dominant contribution arising from the additional term (15.29) of the current whereas the terms which arise from the change in the wave functions by the external perturbation are to be neglected in this case. The first terms give, according to (15.7), a contribution

$$(\vec{\Phi}_{tr})'_{n,m} = \frac{e}{mc} \vec{\varphi}_{tr}^+ e^{i(p_{n,m} + \nu)t} \sum_{a=1}^N \int e^{-i(\vec{K}_x^{(a)} + i(\vec{K}'_x^{(a)}))} u_n^* u_m d^3 x^{(1)} \dots d^3 x^{(N)}, \quad (15.33)$$

to the vector potential of the scattered radiation. Here K and K' are the wave vectors of the incoming and scattered light-waves.

$$\vec{K} = \frac{\nu}{c} \vec{n}; \quad \vec{K}' = \frac{\nu'}{c} \vec{n}'.$$

If we consider somewhat generally the incident light of frequency ν , which is composed of plane waves along various directions with vector potential given by

$$\Phi_k = \Phi_k^+(x_1, x_2, x_3) e^{i\nu t} + \Phi_k^-(x_1, x_2, x_3) e^{-i\nu t}$$

we obtain, instead of (15.33)

$$(\vec{\Phi}_{tr})_{n,m} = \frac{e}{mc} e^{i(p_{n,m} + \nu)t} \sum_{a=1}^N \int \vec{\Phi}_{tr}^+(x_1^{(a)}, x_2^{(a)}, x_3^{(a)}) e^{i(\vec{K}_x^{(a)})} u_n^* u_m d^3 x^{(1)} \dots d^3 x^{(N)}.$$

On account of the completeness relation, the total light intensity of all frequencies ν' scattered along a particular direction can then be written, replacing the various values of K' by a single mean value, as

$$S = \frac{1}{4\pi} \frac{\nu'^2}{m^2 c^2} \frac{e^2}{m^2 c^2} 2 \int \left| \sum_{a=1}^N \vec{\Phi}_{tr}^+(x_1^{(a)}, x_2^{(a)}, x_3^{(a)}) e^{i(\vec{K}_x^{(a)})} \right|^2 u_n^* u_n d^3 x^{(1)} \dots d^3 x^{(N)}. \quad (15.34)$$

Therefore, for an incident plane wave, we have

$$S = \frac{1}{4\pi} \frac{\nu'^2}{m^2 c^2} \frac{e^2}{m^2 c^2} 2 \int \left| \sum_{a=1}^N e^{i(-\vec{K} + \vec{K}') \vec{x}^{(a)}} \right|^2 u_n^* u_n d^3 x^{(1)} \dots d^3 x^{(N)}. \quad (15.34')$$

Since in this expression only the density $u_n^* u_n$ in the initial state n appears and not the densities of the other states, it is, in principle, possible, within the region of validity of this formula, e.g., by using convergent light the intensity of which at a given position is greater than that at another position, to measure the density distribution of the particles in this state. Similarly, it is possible by investigating the spectral intensity distribution of the scattered light in the case of incoming plane waves and using (15.33), to measure the momentum distribution of a bound particle in the initial state. This point was already discussed in Secs. 2 and 11. The validity of

* The reader is referred to G. Källén, Quantum Electrodynamics, George Allen & Unwin Ltd., London and Springer Verlag, New York (1972).

the formula given and the possibility of a simple and direct determination of the density of a particle in coordinate and momentum spaces by investigating the scattered radiation is, however, limited due to the relativistic corrections which have been neglected. As soon as the frequency of the scattered radiation becomes comparable to mc^2/h , the density and current distributions of a stationary state cannot be directly determined for various reasons.

In the above, only the emission and scattering of light were considered but not the associated changes in the stationary states of the atom. But a complete theory should also account for the temporal increase in the probability that the atom, on emission, is to be found in a less excited state. In order to know how far this is possible, let us investigate again the influence of an incident plane wave on the atom on the basis of the perturbation function (15.28). But we shall seek in this case the time-dependent solution which coincides, at $t = 0$, with the unperturbed solution. This means that we set, for the perturbed eigenfunction

$$\psi = \sum_n c_n(t) e^{-i \frac{E_n}{\hbar} t} u_n$$

and make the expansion

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

where $c_n^{(0)}$ is independent of time, $c_n^{(1)}$ is linear and $c_n^{(2)}$ quadratic in the amplitude of the incoming wave, etc. Further $c_n^{(1)}, c_n^{(2)}, \dots$ vanish at $t = 0$ (cf. Sec. 10). Then we can write

$$c_m^{(1)} = i \sum_n T_{m,n} c_n^{(0)}, \quad (15.35)$$

where the Hermitian matrix T will be, as the calculation shows, of the form

$$T_{m,n} = \frac{e^{i(-v_{n,m}+\nu)t} - 1}{(-v_{n,m} + \nu)} \vec{V}_{n,m} \vec{\mathcal{E}}^{(+)} + \frac{e^{-i(v_{n,m}+\nu)t} - 1}{(v_{n,m} + \nu)} \vec{V}_{n,m}^* \vec{\mathcal{E}}^{(-)} \quad (15.36)$$

for an incident plane wave with the field strengths

$$\vec{\mathcal{E}} = \vec{\mathcal{E}}^{(+)} e^{i(\nu t - \vec{k} \cdot \vec{x})} + \vec{\mathcal{E}}^{(-)} e^{-i(\nu t - \vec{k} \cdot \vec{x})}.$$

The matrix $\vec{V}_{n,m}$ here is not necessarily Hermitian. Of special interest is the behaviour of the solution at the resonance, i.e. at the point ν , where one of the two denominators in (15.36) vanishes ($\nu = -v_{n,m} = v_{n,m}$ and $\bar{\nu} = v_{n,m}$). This gives rise in $|c_m^{(1)}(t)|^2$ to terms which increase linearly with time when they are summed over a small interval of different values of ν . These terms are (the others being neglected)

$$\begin{aligned} |c_m^{(1)}(t)|^2 &= c_m^{*(1)}(t) c_m^{(1)}(t) = \sum_n \left| \frac{e^{i(-v_{n,m}+\nu)t} - 1}{-v_{n,m} + \nu} \right|^2 (\vec{V}_{m,n} \vec{\mathcal{E}}^{(-)}) (\vec{V}_{m,n}^* \vec{\mathcal{E}}^{(+)}), \\ &\quad + \sum_n \left| \frac{e^{-i(v_{n,m}+\nu)t} - 1}{v_{n,m} + \nu} \right|^2 (\vec{V}_{n,m}^* \vec{\mathcal{E}}^{(+)})(\vec{V}_{n,m} \vec{\mathcal{E}}^{(-)}). \end{aligned} \quad (15.37)$$

For the resonance position $\nu = v_{n,m}$ (final state energy smaller than that of the initial state) we get after summation over ν

$$|c_m^{(1)}(t)|^2 = t \cdot B_m^* \varrho_r,$$

where B_m^* still depends on the direction and polarisation of the incident radiation.

Similarly for $\nu = \nu_{m,n}$ (final state energy higher than that of the initial state)

$$|c_m^{(1)}(t)|^2 = t \cdot B_m^m c_m,$$

where p_v is the density of the incident radiation. The first case corresponds to induced emission and the second to absorption. *Spontaneous emission is not accounted for here.* In order to get it, we must introduce a new and seemingly arbitrary, prescription analogous to the prescription I. *Prescription II.* $|c_m^{(1)}(t)|^2$ is written formally, in the order $c_m^{*(1)}c_m^{(1)}$ with due consideration for the order of the factors $\mathcal{E}^{(+)}$ and $\mathcal{E}^{(-)}$. No extra term is to be added whenever $\mathcal{E}^{(+)}$ stands before $\mathcal{E}^{(-)}$ [cf. eq. (15.37)], while wherever $\mathcal{E}^{(-)}$ stands before $\mathcal{E}^{(+)}$ we must write $p_v + \frac{2\hbar\nu^3}{c^3}$ instead of p_v .

The justification of Prescriptions I and II introduced ad hoc in this chapter is to be found only in Dirac's quantisation of the radiation field. On the other hand, Prescription I is sufficient in order to discuss correctly the interference experiments and the problems connected with coherence of radiation. This will be demonstrated in the following section.

16. Application of the Semiclassical Theory to the Coherence Properties of Radiation¹

We consider the spontaneous emission of light and shall now investigate when the light emitted from two identical atoms is coherent. Let the initial state be described by the coefficients c_n and c_m of the expansion of the wave functions of the atoms in terms of their eigenfunctions. The matrix elements of the total electric field strength at the space-point P are then of the form

$$\vec{E}_{n,n';m,m'} = \delta_{n',m'} a_{n,m} e^{i\nu_{n,m}(t - \frac{R_P}{c})} + \delta_{n,m} a_{n',m'} e^{i\nu_{n',m'}(t - \frac{R'_P}{c})}. \quad (16.1)$$

Here n and m are the quantum numbers of the pairs of states considered, for one of the atoms and n' and m' , those for the other. The matrix elements of the field strength of the light emitted from the first (second) atom are diagonal with respect to the quantum numbers of the second (first) atom. If these atoms are identical, then for corresponding transitions $a_{n,m} = a_{n',m'}$; $\nu_{n,m} = \nu_{n',m'}$. Further R_P and R'_P represent the distances from the field point P of the atoms, which are considered fixed. For the expectation value of the square of the electrical field strength at the point P , modified according to the Prescription I, we then obtain:

$$\begin{aligned} (\vec{E}^{(+)} \vec{E}^{(-)}) &= \sum_{n,n';m,m'} c_n^* c_{n'}^{*} \mathcal{E}_{n,n';l,l'}^{(+)} \mathcal{E}_{l,l';m,m'}^{(-)} c_m c_{m'} \\ &= \sum_{E_l < E_n} c_n^* a_{n,l} a_{l,m} c_m e^{i\nu_{n,m} t} + \sum_{E_{l'} < E_{m'}} c_{n'}^* a_{n',l'} a_{l',m'} c_{m'} e^{i\nu_{n',m'} t} + \\ &\quad + \sum_{\substack{E_m < E_n \\ E_{n'} < E_{m'}}} c_n^* c_{n'}^* a_{n,m} a_{n',m'} c_m c_{m'} e^{i[(\nu_{n,m} + \nu_{n',m'}) t - \frac{1}{c}(\nu_{n,m} R_P + \nu_{n',m'} R'_P)]} + \\ &\quad + \sum_{\substack{E_m < E_n \\ E_{n'} < E_{m'}}} c_n^* c_{m'}^* a_{m,n} a_{m',n'} c_n c_{n'} e^{i[(\nu_{m,n} + \nu_{m',n'}) t - \frac{1}{c}(\nu_{m,n} R_P + \nu_{m',n'} R'_P)]}. \end{aligned}$$

¹ Here we make only some remarks of a general nature.

The first two terms correspond respectively to the presence of the first and the second atom alone. We are interested in the last two terms which are the interference terms. Remembering that $a_{n,m}$ and $a_{n',m'}$ are Hermitian matrices of which the diagonal elements vanish, the form of the time average becomes very simple if we do not worry about degeneracy here. In the first two sums only the terms with $m = n$, $m' = n'$ give non-vanishing contributions, whereas in the last two sums, the terms with $m' = n$; $n' = m$ contribute. The intensity of light of frequency $\nu_{n,m}$ at the point P will then be

$$J(\nu_{n,m}) = |a_{n,m}|^2 \left\{ |c_n|^2 + |c'_n|^2 + c_n^* c_m c'_n c_m^* e^{-i \frac{\nu_{n,m}}{c} (R_P - R'_P)} + c_n c_m^* c'_n^* c'_m e^{+i \frac{\nu_{n,m}}{c} (R_P - R'_P)} \right\}.$$

We can simplify this expression still further by introducing the geometrical path difference

$$\Delta = \frac{\nu_{n,m}}{c} (R_P - R'_P)$$

and the phases of the probability amplitudes of the atoms according to

$$\begin{aligned} c_n &= |c_n| e^{i\delta_n}; & c_m &= |c_m| e^{i\delta_m}; & \delta_{n,m} &= \delta_n - \delta_m, \\ c'_n &= |c'_n| e^{i\delta'_n}; & c'_m &= |c'_m| e^{i\delta'_m}; & \delta'_{n,m} &= \delta'_n - \delta'_m. \end{aligned}$$

Then

$$J(\nu_{n,m}) = |a_{n,m}|^2 \{ |c_n|^2 + |c'_n|^2 + 2|c_n||c_m||c'_n||c'_m| \cos(\delta_{n,m} - \delta'_{n',m'} + \Delta) \}. \quad (16.2)$$

From this it is immediately seen that no interference can occur, if initially, one of the two atoms was present only in an excited state ($c_m = 0$ or $c'_m = 0$); further, the phases δ_n of the individual eigenfunctions of the atoms can never be observed. A wave-packet can be constructed from the ground state and an excited state in both of the atoms with a definite phase relationship, $\delta_{n,m} - \delta'_{n',m'}$ by exciting both atoms with the same light. In this sense, therefore, resonance radiation is coherent.

It is instructive to discuss briefly also the modifications that arise if we consider the atoms to be moving freely. We then have to introduce besides the states n, m, \dots of the electrons in the atom, also the co-ordinates Q of its centre of gravity, so that the probability amplitudes c_n are now functions of Q . To decide as to how the matrix elements of the field strengths of the emitted light will be modified, let us go back to the expression (15.2) of the classical theory. In order to effect the retardation contained in (15.4) through plane waves, we must assume that the distance of the point P from the atom is large compared to the dimensions of the wave-packet described by $c_n(Q)$ (i.e. compared to the inaccuracy in the definition of the position of the centre of gravity of the atom), which can always safely be done. Further, the contributions of the currents of the atomic nuclei to the light emissions are negligibly small. In the expressions (15.4) for the retarded current of the electrons, the co-ordinates of the centre of mass (C.M.) remain unaltered after integration with respect to the relative co-ordinates of the particles. This is because of the retardation in \vec{x}_0 which represents the sum of the relative and centre of mass co-ordinates. Thus, only the factor

$$e^{i \nu_{n,m} \frac{1}{c} (\vec{Q}_n)} = e^{i (\vec{K}_{n,m} \vec{Q})}$$

(with $\vec{K}_{n,m} = \frac{\nu_{n,m}}{c} \vec{n}$) remains in the matrix element of the emitted field strength taken

between the stationary states (n, m) of the electron system with time dependence given by the factor $e^{i\bar{H}_{n,m}t}$. The co-ordinates of the centre of mass do not enter here in any other way. In the place of the matrix element $a_{n,m}$ in n -space we now have, therefore, the matrix element

$$a_{n,m}(Q, Q') = a_{n,m} e^{i(\vec{K}_{n,m}\vec{\Phi})} \delta(Q - Q') \quad (16.3)$$

in the (n, Q) -space. It is thus instructive to go over to the momentum space of the atom by means of the transform:

$$c_n(P) = \int c_n(Q) e^{-\frac{i}{\hbar}(\vec{P}\vec{\Phi})} dQ$$

[cf. eq. (3.4)]. In eq. (16.3)

$$a_{n,m}(P, P') = a_{n,m} \delta(-\vec{P} + \vec{P}' + \hbar \vec{K}_{n,m}). \quad (16.3')$$

This expression signifies that there is a recoil connected with the emission of light which corresponds exactly to the conservation of momentum, if we associate with the emitted light a momentum of magnitude $\hbar v/c$ and direction along its propagation, in agreement with the well-known result of Einstein* derived from the concept of the photon. The recoil is in principle always observable if the extension of the wave-packet $c_n(P)$ of the initial state in momentum space is small compared to $\hbar v/c$. The recoil must also show itself in Doppler-effect of the emitted light but we shall neglect it here. (This is consistent, since the radiation damping has also been neglected.) Then the total intensity of the light emitted by an atom in the initial state $c_n(Q)$ along the direction \vec{n} is given by

$$\text{if } J(v_{n,m}) = |a_{n,m}|^2 |c_n|^2,$$

$$|c_n|^2 = \int |c_n(Q)|^2 dQ = \int |c_n(P)|^2 dP. \quad (16.4)$$

The application of the above ideas to the intensity of the light emitted by two identical atoms is given, on using the abbreviation

$$C_{n,m} = |C_{n,m}| e^{-i\delta_{n,m}} = \int c_n^*(Q) e^{i(\vec{K}_{n,m}\vec{\Phi})} c_m(Q) dQ = C_{m,n}^* \quad (16.5)$$

and a similar one for $C'_{n,m}$, by the expression

$$J(v_{n,m}) = |a_{n,m}|^2 \{ |c_n|^2 + |c'_n|^2 + 2 |C_{n,m}| |C'_{n,m}| \cos(\delta_{n,m} - \delta'_{n,m} + \Delta) \}, \quad (16.6)$$

instead of (16.2), if in $\Delta = \frac{v_{n,m}}{c}(R_p - R'_p)$, R and R' are counted from a fixed point [the same as Q in (16.5) and Q' in the similar definition of $C'_{n,m}$]. The meaning of the results obtained above is that for coherence of resonance radiation it is not only necessary that in the initial state both the atoms are present in the ground state or in the excited state but also that these states can occur with non-vanishing probability for the same position of the centre of mass of the atoms. If the functions $c_n(Q)$ and $c_m(Q)$ do not overlap, as is the case, if both the states are completely separated by an external field, then the product $c_n^*(Q) c_m(Q)$ vanishes everywhere and hence also the interference term in (16.6). This is an example of the fact that every arrangement which permits us to determine the state in which the atom is to be found

* A. Einstein, Ann. d. Physik 17, 132 (1905).

removes the possibility of the interference of the light emitted by this atom with the light emitted by other atoms.²

Finally we shall examine the question of coherence of the radiation emitted by an atom in different directions, say in the directions, \vec{n}_1 and \vec{n}_2 , since this question concerns the historical controversy between "needle radiation" and "spherical waves". All devices for testing the possible interference of radiation along these directions amount to reuniting the two pencils of light at a point P after suitable reflections and refractions. At this point, therefore, the classical field strengths will be a linear combination of $\mathcal{E}(\vec{n}_1)$ and $\mathcal{E}(\vec{n}_2)$, i.e. the field strengths of the bundles originally emitted in the directions \vec{n}_1 and \vec{n}_2 . If J_0 is the sum of the intensities of the bundles, emitted in these directions, as observed at the point P when there is no interference, then for the true intensity we have, classically

$$J = J_0 + \text{const } \mathcal{E}(\vec{n}_1) \mathcal{E}(\vec{n}_2).$$

We have, therefore, to calculate quantum-theoretically the expectation value of

$$\mathcal{E}^{(+)}(\vec{n}_1) \mathcal{E}^{(-)}(\vec{n}_2) + \mathcal{E}^{(+)}(\vec{n}_2) \mathcal{E}^{(-)}(\vec{n}_1)$$

as a measure of the coherence of the pencils. This expectation value is proportional to

$$D = \int dQ c_n^*(Q) \int a_{n,m}(\vec{n}_1; Q, Q'') a_{m,n}(\vec{n}_2; Q'', Q') dQ'' c_n(Q') dQ' + \dots$$

or also to

$$D = \int dP c_n^*(P) \int a_{n,m}(\vec{n}_1; P, P'') a_{m,n}(\vec{n}_2; P'', P') dP'' c_n(P') dP' + \dots,$$

where $+ \dots$ denotes the term obtained from interchanging \vec{n}_1 and \vec{n}_2 in the first term. On the basis of (16.3), (16.3'), we obtain

$$D = 2 \int |c_n(Q)|^2 \cos \frac{\hbar v_{n,m}}{c} (\vec{n}_2 - \vec{n}_1) \cdot \vec{Q} dQ \quad (16.7)$$

or also

$$D = \int \left\{ c_n^*(P) c_n \left(P + \frac{\hbar v_{n,m}}{c} (n_2 - n_1) \right) + c_n(P) c_n^* \left(P + \frac{\hbar v_{n,m}}{c} (n_2 - n_1) \right) \right\} dP. \quad (16.7')$$

From these expressions we see that the possibility of determining by a recoil measurement, whether the photon is emitted in the direction \vec{n}_1 or \vec{n}_2 stands in an exclusive relationship with an arrangement which enables us to test the interference between the light bundles emitted in the directions \vec{n}_1 and \vec{n}_2 . A measurement of the recoil of the required sort is possible only if initially the momentum of the particle is defined more accurately than $\frac{\hbar v_{n,m}}{c} |n_2 - n_1|$. Then, however, $c_n(P)$ can be different from zero only in a region ΔP which is smaller than $\frac{\hbar v_{n,m}}{c} |n_2 - n_1|$. In this case D vanishes as is clear from (16.7'). On the other hand, in order to define the path difference between the bundles emitted along \vec{n}_1 and \vec{n}_2 with precision, $c_n(Q)$ can be different from zero only in region ΔQ which is small compared to $\frac{c}{v_{n,m}} \frac{1}{|n_2 - n_1|}$. The fact that these two requirements are mutually contradictory is a direct consequence of Heisenberg's uncertainty relation which, in turn, is satisfied in the calculation of

² Cf. for this W. Heisenberg, Z. Physik 43, 172 (1927). At that time the question of the connection between the phases of the eigenfunctions of the atom and the properties of the emitted light was not yet clarified.

$c_n(P)$ from $c_n(Q)$.

The coherence of the radiation *scattered* from the atoms can be discussed in the same way as was done here for the simplest cases of light emission. But it must be explicitly stressed that the treatment given above is still incomplete since radiation damping is not taken into account. This can be done only by means of Dirac's theory of the photon.

CHAPTER IX

The Relativistic One-Particle Problem

17. Introduction

In contrast to non-relativistic quantum mechanics, which can be considered logically complete, we possess today only fragments of a relativistic quantum mechanics. We shall treat here only the relativistic one-particle problem which is dominated by Dirac's wave equation¹ and which describes the behaviour of an electron in a given external electromagnetic field.

The problem of field quantisation is reserved for the next chapter.

18. Dirac's Wave Equation for the Electron. Free Particle Case

The fundamental relationship between momentum and energy on the one hand and the wave vector and frequency of the waves on the other, given by the relation (I) of Section 1,

$$\vec{p} = \hbar \vec{k}; \quad E = \hbar \omega$$

is relativistically invariant. For $(\vec{p}, i \frac{E}{c})$, $(\vec{k}, i \frac{\omega}{c})$ are both four-vectors which, therefore, transform under a Lorentz transformation in the same way. Hence it is natural to retain these relations as the basis of a relativistic quantum theory. In classical relativistic mechanics, we have the relation given by eq. (1.5) of Sec. 1,

$$\frac{E^2}{c^2} = m^2 c^2 + \sum_{i=1}^3 p_i^2, \quad (18.1)$$

between the energy and momentum of a particle with a rest mass m . Substituting (I) in (18.1) we get

$$\frac{\omega^2}{c^2} = \frac{m^2 c^2}{\hbar^2} + \sum_i k_i^2 = \frac{\omega_0^2}{c^2} + \sum_i k_i^2 \quad (18.2)$$

with

$$\omega_0 = \frac{mc^2}{\hbar}. \quad (18.3)$$

The most general superposition of plane waves

$$\psi(\vec{x}; t) = \int A(\vec{k}) e^{i(\vec{k} \cdot \vec{x} - \omega t)} d^3 k, \quad (18.4)$$

¹ P.A.M. Dirac, Proc. Roy. Soc. Lond., 117, 610; 118, 341 (1928).

in which ω and \vec{k} are connected by (18.2), satisfies the differential equation*

$$\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = D\psi - \frac{m^2 c^2}{\hbar^2} \psi, \quad (18.5)$$

and conversely (18.4) is in essence the most general solution of (18.5). Here it is to be remembered that according to (18.2) there are two values of ω belonging to a given value of \vec{k} , one positive and the other negative, namely

$$\frac{\omega}{c} = + \sqrt{\frac{m^2 c^2}{\hbar^2} + \sum_i k_i^2} \quad \text{and} \quad \frac{\omega}{c} = - \sqrt{\frac{m^2 c^2}{\hbar^2} + \sum_i k_i^2}. \quad (18.2')$$

In general both values of ω can occur in eq. (18.4). The wave equation (18.5) is relativistically invariant, if ψ is treated as a scalar (under a Lorentz transformation).

Till now only the fundamental de Broglie relations (I) and the wave-theoretical superposition principle have been used. If we follow further the corresponding development of non-relativistic quantum mechanics in a logical fashion, we see that the next step consists in the introduction of a probability density $W(x_1, x_2, x_3)$ which states how probable it is that the particle is found at time t in the spatial region $(x_1, x_1 + dx_1 \dots x_3, x_3 + dx_3)$. In case such a probability density $W(x)$ exists, it must be, first of all, positive (or zero):

$$W(x) \geq 0 \quad (18.6)$$

and secondly it must satisfy the condition

$$\frac{d}{dt} \int W(x) d^3x = 0 \quad (18.7)$$

as a consequence of the wave equation. This is necessary in order that it may be normalised according to

$$\int W(x) d^3x = 1. \quad (18.7')$$

We also require the invariance of this normalisation under Lorentz-transformations:

$$\int W(x) d^3x = \text{Invariant.} \quad (18.8)$$

If we now try to construct from eq. (18.5) an expression which satisfies the conditions (18.7) and (18.8), we are necessarily led to

$$\varrho(x) \equiv W(x) = -\psi^* \frac{1}{c} \frac{\partial \psi}{\partial t} + \psi \frac{1}{c} \frac{\partial \psi^*}{\partial t}.$$

For, with the introduction of the vectors

$$\vec{i} = c(\psi^* \text{grad } \psi - \psi \text{grad } \psi^*), \quad (18.9)$$

where ψ^* is the complex conjugate of ψ , the continuity equation

$$\frac{\partial \varrho}{\partial t} + \text{div } \vec{i} = 0$$

will be satisfied and we can combine ϱ and \vec{i} into a four-vector s with the components

$$s_\nu = \left(\frac{1}{c} \vec{i}, i \varrho \right),$$

*This equation is called the Klein-Gordon equation which in its second-quantised version is applicable to spin-zero particles like the π and K mesons.

according to

$$s_\nu = \psi^* \frac{\partial \psi}{\partial x_\nu} - \psi \frac{\partial \psi^*}{\partial x_\nu}. \quad (18.10)$$

From the above, eqs. (18.7) and (18.8) follow.² This corresponds to the attempts made originally by many authors.³ Since according to (18.5) both ψ and $\partial \psi / \partial t$ can be chosen arbitrarily at a particular moment eq. (18.6) is violated, and hence the expression (18.10) for the four-current is not physically permissible.⁴

On the basis of the above result we may doubt whether in relativistic quantum mechanics the probability density of a particle is a meaningful concept. For, (1) its definition contains already a preference of time before space, in that the space co-ordinates x_k are allowed to have a spread dx_k whereas the time is fixed exactly; (2) the determination of the position of the particle is no longer possible by a *direct* measurement, in case we are considering regions which are small compared to the wave-length of the matter waves and simultaneously velocities of the particles which are comparable with the velocity of light [cf. (2.4) and (2.5)]; (3) there does not, in fact, exist in the case of photons a probability density satisfying the requirements (18.6), (18.7) and (18.8), as we shall see later.

In spite of this Dirac was able to show that an ansatz for $W(x)$ satisfying the requirements (18.6) to (18.8) is possible if we introduce *multi-component* ψ -functions ψ_ρ , $\rho = 1, 2, \dots$ all of which satisfy the eq. (18.5) in the free-particle case. In view of the great success of this ansatz, in that it automatically explains the spin of electrically charged particles, it would be advisable to follow up all the consequences of this ansatz and to take up a discussion of the objections mentioned above, only later, when we actually encounter the principal difficulties (states of negative energy) of Dirac's theory.

Dirac's ansatz consists in retaining the expression

$$\varrho = W(x) = \sum_\sigma \psi_\sigma^* \psi_\sigma \quad (18.11)$$

which alone guarantees the positive definite character of $W(x)$. Since we require

$$\frac{d}{dt} \int \sum_\sigma \psi_\sigma^* \psi_\sigma d^3x = \int \sum_\sigma \left(\frac{\partial \psi_\sigma^*}{\partial t} \psi_\sigma + \psi_\sigma^* \frac{\partial \psi_\sigma}{\partial t} \right) d^3x = 0, \quad (18.12)$$

$\partial \psi_\sigma / \partial t$ and $\partial \psi_\sigma^* / \partial t$ cannot be arbitrary at a fixed moment of time; the ψ_σ must,

² In the following we denote by Greek letters indices running from 1 to 4, by Latin letters those running from 1 to 3, by x_4 the imaginary time co-ordinate $x_4 = ict$ and by x_0 the real time co-ordinate $x_0 = ct$, so that $x_4 = ix_0$.

³ E. Schrödinger, Ann. d. Phys. 81, 129 (1926), in particular Sec. 6. – O. Klein, Z. Physik 37, 895 (1926) – V. Fock, Z. Physik, 38, 242; 39, 226 (1926). – J. Kudar, Ann. d. Phys. 81, 632 (1926). Concerning the expressions for the four-current, see W. Gordon, Z. Physik 40, 117 (1926). All these authors consider directly the general case of a charged particle in an external electromagnetic field which will be dealt with in Sec. 21.

⁴ If we had not used the requirement (18.8) but only the requirement (18.7), then the ansatz

$$\varrho = \frac{1}{2} \left[\left(\frac{\partial \psi}{\partial t} \right)^2 + (\text{grad } \psi)^2 + \frac{m^2 c^2}{\hbar^2} \psi^2 \right]$$

would have been possible, since then from (18.5) would follow:

$$\frac{\partial \varrho}{\partial t} + \text{div} \left(\frac{\partial \psi}{\partial t} \text{grad } \psi \right) = 0$$

This ansatz is remarkable because we obtain it with a single *real* function ψ and with it condition (18.6) is fulfilled. But then, $\int \varrho d^3x$ appears as the fourth-component of a vector instead of as a scalar.

therefore, satisfy differential equations of the first order in $\partial/\partial t$. In order to satisfy the relativistic invariance of the equations subsequently it is necessary to require the differential equations to be of the first order in the spatial derivatives $\partial/\partial x_k$ also. Following Dirac, we, therefore, start from the differential equations

$$\frac{1}{c} \frac{\partial \psi_\rho}{\partial t} + \sum_{\sigma} \left(\sum_{k=1}^3 \alpha_{\rho\sigma}^k \frac{\partial \psi_\sigma}{\partial x_k} + i \frac{mc}{\hbar} \beta_{\rho\sigma} \psi_\sigma \right) = 0. \quad (18.13)$$

The question of the values that the numerical coefficients $\alpha_{\rho\sigma}^k$ and $\beta_{\rho\sigma}$ can take and the number of values which each of the indices ρ and σ can take are left open for the time being. In order that (18.12) follows from (18.13), it is sufficient to assume that

$$\alpha_{\rho\sigma}^{*\rho} = \alpha_{\sigma\rho}^\rho; \quad \beta_{\rho\sigma}^* = \beta_{\sigma\rho}. \quad (18.14)$$

Then from (18.13) follows

$$\frac{1}{c} \frac{\partial \psi_\rho^*}{\partial t} + \sum_{\sigma} \left(\sum_{k=1}^3 \frac{\partial \psi_\sigma^*}{\partial x_k} \alpha_{\rho\sigma}^k - i \frac{mc}{\hbar} \psi_\rho^* \beta_{\rho\sigma} \right) = 0. \quad (18.13^*)$$

If we multiply (18.13) by ψ_ρ^* and sum over ρ , multiply (18.13*) by ψ_σ and sum over σ , then we obtain the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{i} = 0, \quad (18.15)$$

if we set

$$\vec{i} = c \sum_{\rho} \sum_{\sigma} \psi_\rho^* \vec{\alpha}_{\rho\sigma} \psi_\sigma. \quad (18.16)$$

To simplify the notation – i.e., in order to avoid the writing of the indices – it is useful to introduce the notation of matrix calculus. Here α^k and β appear as square Hermitian matrices – their Hermiticity is equivalent to the requirement (18.14) – while ψ is to be understood as a column vector (with components ψ_ρ) and ψ^* as a row vector (with components ψ_ρ^*). In order to obtain meaningful expressions from the rule for matrix multiplication, ψ^* must always stand to the *left* of the matrices α^k and β and ψ always to their *right*. Then eqs. (18.13), (18.13*), (18.11), and (18.16) can simply be rewritten as

$$\frac{1}{c} \frac{\partial \psi}{\partial t} + \sum_{k=1}^3 \alpha^k \frac{\partial \psi}{\partial x_k} + i \frac{mc}{\hbar} \beta \psi = 0, \quad (18.13')$$

$$\frac{1}{c} \frac{\partial \psi^*}{\partial t} + \sum_{k=1}^3 \frac{\partial \psi^*}{\partial x_k} \alpha^k - i \frac{mc}{\hbar} \psi^* \beta = 0, \quad (18.13'')$$

$$\rho = (\psi^* \psi), \quad (18.11')$$

$$\vec{i} = c (\psi^* \vec{\alpha} \psi). \quad (18.16')$$

Just as the wave equation (second order) for each of the field strengths follows from Maxwell's equations (first order) for the field strengths, eq. (18.5), namely,

$$-\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} + \sum_{k=1}^3 \frac{\partial^2 \psi}{\partial x_k^2} - \frac{m^2 c^2}{\hbar^2} \psi = 0$$

for each of the components of ψ should now follow from eq. (18.13). To test this, let us apply, from the left, the operator

$$-\frac{1}{c} \frac{\partial}{\partial t} + \sum_i \alpha^i \frac{\partial}{\partial x^i} + i \frac{mc}{\hbar} \beta$$

on eq. (18.13). This is the only operation which makes the terms of the *first* order in $\frac{\partial}{\partial t}$ to drop out. We get immediately

$$-\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} + \sum_l \sum_k \left\{ \alpha^l \alpha^k \frac{\partial^2 \psi}{\partial x_l \partial x_k} \right\} + i \frac{mc}{\hbar} \sum_k (\alpha^k \beta + \beta \alpha^k) \frac{\partial \psi}{\partial x_k} - \frac{m^2 c^2}{\hbar^2} \beta^2 \psi = 0,$$

or if we symmetrise the term in l and k ,

$$-\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} + \sum_l \sum_k \frac{1}{2} (\alpha^k \alpha^l + \alpha^l \alpha^k) \frac{\partial^2 \psi}{\partial x_k \partial x_l} + i \frac{mc}{\hbar} \sum_k (\alpha^k \beta + \beta \alpha^k) \frac{\partial \psi}{\partial x_k} - \frac{m^2 c^2}{\hbar^2} \beta^2 \psi = 0.$$

Comparison with (18.5) shows that for its validity, it is necessary and sufficient to require that

$$\frac{1}{2} (\alpha^k \alpha^l + \alpha^l \alpha^k) = \delta_{l,k} I; \quad \alpha^k \beta + \beta \alpha^k = 0; \quad \beta^2 = I, \quad (I)$$

(where I is the unit matrix). The first set of relations is equivalent to

$$(\alpha^k)^2 = I; \quad \alpha^k \alpha^l = -\alpha^l \alpha^k \text{ for } k \neq l. \quad (I')$$

It is to be noted that the four matrices α^k and β form an anticommuting set* and β enjoys the same properties as α^k .

We have still to discuss, if and in how many ways, the relations (I) and the Hermiticity requirements can be fulfilled. It turns out that the number of rows of the matrices satisfying relation (I) must be at least four. A possible solution with matrices having four rows is obtained on using the two-by-two matrices appearing in the non-relativistic theory of spin:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (18.17)$$

which satisfy the relations

$$\left. \begin{aligned} \sigma_1 \sigma_2 &= -\sigma_2 \sigma_1 = i \sigma_3, \dots, \\ \sigma_1^2 &= I, \dots \end{aligned} \right\} \quad (18.18)$$

(The dots in (18.18) indicate that the remaining relations are to be obtained by cyclic permutations of the indices.) A possible solution of (I) in which β is diagonal in addition, is then given by

$$\alpha^k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}; \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}. \quad (18.19)$$

Here the four matrices α^k and β have been written in the partitioned notation of 2×2 matrices; I is the 2×2 unit matrix, 0 is the two-by-two null matrix. The given matrices further satisfy the Hermiticity requirement.

As regards the existence of other solutions of eq. (I), we may state that, in any case, the set of transformations

$$\alpha'^k = S \alpha^k S^{-1}; \quad \beta' = S \beta S^{-1} \quad (18.20)$$

with a *unitary* (in order to preserve Hermiticity) but otherwise arbitrary matrix S is

*But these four matrices do not form the maximal set of anticommuting matrices; see eq. (19.12).

possible without altering the validity of the relations (I). Further, a trivial extension of the given matrices α^k, β to matrices of higher dimensions of the form

$$A_k = \begin{pmatrix} \alpha^k, 0, 0 \dots \\ 0, \alpha^k, 0 \dots \\ 0, 0, \alpha^k \dots \\ \dots \dots \end{pmatrix}; \quad B = \begin{pmatrix} \beta, 0, 0, 0 \dots \\ 0, \beta, 0, 0 \dots \\ 0, 0, \beta, 0 \dots \\ \dots \dots \end{pmatrix}$$

where all the α^k and β are 4×4 matrices is possible. Finally, we can also transform these A_k and B by the "enlarged" unitary matrix S :

$$A'_k = S A_k S^{-1}; \quad B' = S B S^{-1}.$$

The decomposition into submatrices is then no longer evident in the A'_k and B' . It can be shown that besides these trivial extensions of the given solution, none other exists.⁵

We have shown so far that the basic equations (18.13), with the ansatz (18.11') and (18.16') for the density and current satisfy the continuity equation and that the original equations (18.5) follow from (18.11'). We still must show that the eqs. (18.13) are also relativistically invariant and that the density and current combine to form a four-current. From the latter circumstance itself then follows automatically the invariance of the normalisation of ψ_ρ by the volume integral $\int \rho d^3x$, as required by (18.8).

19. Relativistic Invariance

In order to investigate the relativistic invariance of the set of equations (18.13), it is convenient to transform it in such a way that the four co-ordinates $x_\mu (\mu = 1, \dots, 4)$ with $x_4 = ict$ appear on equal footing. To this end we multiply (18.13) on the left by $-i\beta$ and obtain (since $\beta^2 = 1$)

$$\sum_{\mu=1}^4 \gamma^\mu \frac{\partial \psi}{\partial x_\mu} + \frac{mc}{\hbar} \psi = 0, \quad (\text{II})$$

where

$$\gamma^4 = \beta; \quad -i\beta\alpha^k = \gamma^k, \quad (19.1)$$

and, therefore

$$\beta = \gamma^4; \quad \alpha^k = i\gamma^4\gamma^k. \quad (19.1')$$

The matrices γ^μ are Hermitian, like α^k and β , and satisfy as a consequence of (I), similar anticommutation relations:

$$\frac{1}{2}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) = \delta_{\mu\nu} \cdot I. \quad (\text{I}'')$$

⁵ The demonstration consists, as may be briefly indicated here, in the following: The 16 linearly independent elements $\gamma^\mu, \gamma^\mu\gamma^\nu, \gamma^\mu\gamma^\nu\gamma^\rho, \gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma$ (where μ, ν, ρ are supposed to be different from one another) form the basis of a hyper-complex number system which falls in the class of semi-simple systems. The centre, i.e. the element which commutes with all elements has the basis 1, consisting of only one single element. Since, in general, the number of the inequivalent irreducible representations of a semi-simple system coincides with the number of basis elements of the centre, only a single irreducible representation exists in our case. The square of its degree f is equal to the number n of the basis elements, $f^2 = n$; therefore, in our case $f = 4$.

Eq. (18.13*) assumes, on substituting (19.1') and dividing by i , the form

$$\frac{\partial \psi^*}{\partial x_\mu} + \sum_k \frac{\partial \psi^*}{\partial x_k} \gamma^k \gamma^\mu - \frac{mc}{\hbar} \psi^* \gamma^k = 0.$$

If we, therefore, set

$$\psi^\dagger = \psi^* \gamma^4; \quad \psi^* = \psi^\dagger \gamma^4, \quad (19.2)$$

it then follows that

$$\sum_\mu \frac{\partial \psi^\dagger}{\partial x_\mu} \gamma^\mu - \frac{mc}{\hbar} \psi^\dagger = 0. \quad (\text{II}^\dagger)$$

The four-vector s_μ , with components

$$s_\mu = \left(\frac{i}{c}, i \boldsymbol{\theta}_\mu \right) \quad (19.3)$$

assumes, according to (18.11') and (18.16'), the form

$$s_\mu = i \psi^\dagger \gamma^\mu \psi. \quad (19.4)$$

Relations (II) and (II †) define ψ^\dagger only up to a factor, for a given ψ , which then is normalised through (19.2). The form of the Dirac equation given here is useful for studying the relativistic invariance properties, whereas the form used earlier possesses the advantage of making transparent the reality properties of the ψ -function.

We consider now the orthogonal co-ordinate transformations

$$\left. \begin{aligned} x'_\mu &= \sum_\nu a_{\mu\nu} x_\nu, & \sum_\mu a_{\mu\sigma} a_{\mu\sigma} &= \delta_{\sigma\sigma}, \\ x_\nu &= \sum_\mu a_{\mu\nu} x'_\mu, & \sum_\sigma a_{\sigma\mu} a_{\sigma\nu} &= \delta_{\mu\nu}, \end{aligned} \right\} \quad (19.5)$$

and set

$$\psi' = (S \psi), \quad (19.6)$$

where S is a 4×4 matrix yet to be determined. Then from the relation

$$\sum_\mu \gamma^\mu \frac{\partial \psi}{\partial x_\mu} + \frac{mc}{\hbar} \psi = 0$$

we should have

$$\sum_\mu \gamma^\mu \frac{\partial \psi'}{\partial x'_\mu} + \frac{mc}{\hbar} \psi' = 0,$$

or

$$\sum_\mu \sum_\nu \gamma^\mu S a_{\mu\nu} \frac{\partial \psi}{\partial x_\nu} + \frac{mc}{\hbar} S \psi = 0,$$

or

$$\sum_\mu \sum_\nu (S^{-1} \gamma^\mu S) a_{\mu\nu} \frac{\partial \psi}{\partial x_\nu} + \frac{mc}{\hbar} \psi = 0.$$

This last equation is satisfied if

$$\sum_\mu (S^{-1} \gamma^\mu S) a_{\mu\nu} = \gamma^\nu$$

or

$$S^{-1} \gamma^\mu S = \sum_\nu a_{\mu\nu} \gamma^\nu. \quad (\text{A})$$

Further from (II †) follows the validity of

$$\frac{\partial \psi^\dagger}{\partial x'_\mu} \gamma^\mu - \frac{mc}{\hbar} \psi^\dagger = 0,$$

if we set

$$\psi^{\dagger'} = (\psi^{\dagger} S^{-1}). \quad (19.6')$$

Similarly, if (A) holds, the expressions (19.4) for s_{μ} , in fact, form a four-vector and

$$J = (\psi^{\dagger} \psi) \quad (19.7)$$

is invariant. Elementary computations show that

$$M_{\mu\nu} = -M_{\nu\mu} = i\psi^{\dagger} \gamma^{\mu} \gamma^{\nu} \psi \quad (\mu \neq \nu) \quad (19.8)$$

is an antisymmetric tensor of the second rank;

$$K_{\mu\nu\rho} = i\psi^{\dagger} \gamma^{\mu} \gamma^{\nu} \gamma^{\rho} \psi, \quad (\mu \neq \rho \neq \nu \neq \mu) \quad (19.9)$$

is a tensor of the third rank, antisymmetric in all the three indices (pseudovector) and

$$N = i\psi^{\dagger} \gamma^5 \psi \quad (19.10)$$

with

$$\gamma^5 = \gamma^1 \gamma^2 \gamma^3 \gamma^4, \quad (19.11)$$

is a pseudoscalar. The last statement means that under co-ordinate transformations we have

$$N' = N |a_{\mu\nu}|,$$

where the determinant $|a_{\mu\nu}|$ of the orthogonal transformation is +1 for proper rotations and -1 under reflections.¹ The matrix γ^5 is remarkable as it satisfies the relations

$$(\gamma^5)^2 = 1; \quad \gamma^5 \gamma^{\mu} + \gamma^{\mu} \gamma^5 = 0. \quad (19.12)$$

Therefore, there exist five independent 4×4 matrices which satisfy the relations (I).² The matrix S is not unitary because of the imaginary character of the fourth coordinate, since the a_{4k} and a_{k4} for $k = 1, 2, 3$ are pure imaginary and only a_{kl} and a_{44} are real. We must, therefore, set the Hermitian conjugate \tilde{S} of S equal to

$$\tilde{S} = \gamma^4 S^{-1} \gamma^4 \quad \text{or} \quad \tilde{S} \gamma^4 = \gamma^4 S^{-1}, \quad (19.13)$$

instead of equating it to S^{-1} . Then, with

$$\psi^{*\prime} = \psi^* \tilde{S}$$

we have in the new co-ordinate system

$$\psi^{\dagger'} = \psi^{*\prime} \gamma^4.$$

which is similar to (19.2). Now it remains to prove that for every orthogonal transformation (19.5) there exists a matrix S , which is a function of $a_{\mu\nu}$, satisfying (A).

¹ Cf. J.v. Neumann, Z. Physik **48**, 868 (1928).

² Between the quantities $J, s_{\nu}, M_{\mu\nu}, K_{\mu\nu\rho}, N$ there exist various quadratic identities like

$$-\sum_{\nu} (s_{\nu})^2 = J^2 + N^2 = K_{111}^2 + K_{411}^2 + K_{341}^2 + K_{331}^2,$$

$$K_{111} s_4 + K_{411} s_3 + K_{341} s_2 + K_{331} s_1 = 0.$$

Cf. for this V. Fock, Z. Physik **57**, 261 (1929); C.G. Darwin, Proc. Roy. Soc. Lond. **120**, 621 (1928); G.E. Uhlenbeck and O. Laporte, Phys. Rev. **37**, 1380 (1931); W. Pauli, in Zeeman Verhandlungen, pp. 31-43 (1935); Ann. Inst. H. Poincaré **6**, 109 (1936).

Since the matrices $\sum a_{\mu\nu} \gamma^\mu$ satisfy the same relations (I'') as the γ^μ as a consequence of (19.5), the existence of such an S follows from the uniqueness (up to an equivalence) of the irreducible representations of the γ^μ . A second independent proof is obtained on using the group property of the orthogonal transformations, in which one shows that the equations (A) can be satisfied for infinitesimal transformations. Such a transformation is given by

$$x'_\mu = x_\mu + \sum \epsilon_{\mu\nu} x_\nu \quad \text{with} \quad \epsilon_{\mu\nu} = -\epsilon_{\nu\mu}, \quad (19.14)$$

where the antisymmetry of the $\epsilon_{\mu\nu}$ ensures that the orthogonality conditions are satisfied. For S , let us then make the ansatz, linear in the $\epsilon_{\mu\nu}$:

$$S = I + \frac{1}{2} \sum_{\mu\nu} \epsilon_{\mu\nu} T^{\mu\nu} \quad \text{with} \quad T^{\mu\nu} = -T^{\nu\mu}, \quad (19.15)$$

where the $T^{\mu\nu}$ are matrices which are labelled by a pair of indices. Now we substitute (19.15) in (A) and retain only quantities of the first order in $\epsilon_{\mu\nu}$. We obtain, then

$$\gamma^\mu \frac{1}{2} \sum_{\lambda} \epsilon_{\lambda\mu} T^{\lambda\nu} - \frac{1}{2} \sum_{\lambda} \sum_{\nu} \epsilon_{\lambda\mu} T^{\lambda\nu} \gamma^\mu = \sum_{\nu} \epsilon_{\mu\nu} \gamma^\nu = \frac{1}{2} \sum_{\lambda} \sum_{\nu} \epsilon_{\lambda\nu} (\delta_{\lambda\mu} \gamma^\nu - \delta_{\nu\mu} \gamma^\lambda).$$

By comparing the coefficients of $\epsilon_{\lambda\nu}$ which are already made antisymmetric in the indices λ and ν , we obtain

$$\gamma^\mu T^{\lambda\nu} - T^{\lambda\nu} \gamma^\mu = \delta_{\lambda\mu} \gamma^\nu - \delta_{\nu\mu} \gamma^\lambda. \quad (A')$$

This equation can, in fact, be satisfied by 4×4 matrices $T^{\lambda\nu}$ and thus the relativistic invariance of the Dirac equation and the vector character of s_ν are proved at the same time. A solution of (A') which is antisymmetric in λ and ν is

$$T^{\lambda\nu} = \frac{1}{2} \gamma^i \gamma^\nu \quad \text{for} \quad \lambda \neq \nu; \quad (T^{\lambda\nu} = 0 \quad \text{for} \quad \lambda = \nu). \quad (19.16)$$

as can be easily verified using (I').

We have now still to test the condition (19.13) which fixes the reality properties of the matrices. We have

$$\sum_{\mu\nu} \epsilon_{\mu\nu}^* \tilde{T}^{\mu\nu} = - \sum_{\mu\nu} \gamma^k \epsilon_{\mu\nu} T^{\mu\nu} \gamma^k. \quad (19.17)$$

Remembering now that ϵ_{ik} is pure imaginary and ϵ_{ik} is real for i, k taking values from 1 to 3, it follows that

$$\tilde{T}^{ik} = +\gamma^k T^{ik} \gamma^i, \quad \tilde{T}^{ik} = -\gamma^i T^{ik} \gamma^k \quad (i, k = 1, 2, 3). \quad (19.17')$$

Due to the definition (19.16), $\tilde{T}^{\lambda\nu} = -T^{\lambda\nu}$, it is verified easily, on the basis of (I') that (19.17') is in fact fulfilled.

It is easy to check the uniqueness of the solution of (A') and (19.17). The equation (A') allows an additional term in the $T^{\lambda\nu}$ which commutes with all γ^μ . Such a term is necessarily of the form

$$\Delta_{\lambda\nu} \cdot I$$

with ordinary numbers $\Delta_{\lambda\nu}$. The equation (19.17) then requires that

$$\Delta = \sum_{\lambda\nu} \epsilon_{\lambda\nu} \Delta_{\lambda\nu}$$

is pure imaginary. Since up to the first order

$$1 + i|\Delta| = e^{i|\Delta|},$$

we have here to deal with an additional (infinitesimal) phase factor which is common to all the four components of ψ . Such a factor is, in fact, arbitrary. Its normalisation through the ansatz (19.16) corresponds to the stipulation that

$$\text{Spur}(T^{\lambda\mu}) = 0, \quad (19.18)$$

if, as usual, we understand by the spur of a matrix the sum of its diagonal elements. For finite transformations, this has the consequence, that

$$\text{Det}(S) = 1, \quad (19.18')$$

where Det denotes the determinant of the matrix.

We remark further that according to (19.16), the $T^{\lambda\mu}$ commute with the matrix γ^5 defined by (19.11):

$$\gamma^5 T^{\lambda\mu} - T^{\lambda\mu} \gamma^5 = 0. \quad (19.19)$$

This has the consequence that for finite transformations which can be generated by continuous continuation of infinitesimal transformations,

$$S^{-1} \gamma^5 S = \gamma^5$$

It has already been mentioned that

$$N = i\psi^\dagger \gamma^5 \psi$$

is a pseudoscalar quantity which is equivalent to the statement following from (A), that

$$S^{-1} \gamma^5 S = \pm \gamma^5 \quad (19.20)$$

with positive and negative signs for orthogonal transformations of the co-ordinates with the determinants +1 and -1, respectively. The latter are the reflections. If we consider, e.g., the reflection

$$x'_k = -x_k \quad \text{for } k = 1, 2, 3; \quad x'_4 = x_4, \quad (19.21)$$

we have, according to (A), to search for an S , for which

$$\gamma^k S = -S \gamma^k \quad \text{for } k = 1, 2, 3; \quad \gamma^4 S = S \gamma^4.$$

From this follows, [with due consideration given to the additional conditions (19.13) and (19.18')] that

$$S = \gamma^4. \quad (19.21')$$

By means of the solution (19.16) of (A'), it is even possible in the special case of a rotation around a co-ordinate axis – in which, therefore, only two of the co-ordinates change – to give the solution of (A) for a *finite* rotation of the co-ordinate frame.³ If we consider, e.g., first a rotation in the (x_1, x_2) -plane which is given by

$$x'_1 = x_1 \cos \omega - x_2 \sin \omega,$$

$$x'_2 = x_1 \sin \omega + x_2 \cos \omega,$$

then, remembering that here the matrices $S(\omega)$ commute with one another for all values of ω , we have, according to (19.16)

$$\frac{dS}{d\omega} = S T^{12} = S \frac{1}{2} \gamma^1 \gamma^2.$$

³ Cf. P.A.M. Dirac, Quantum Mechanics, loc. cit.

Recalling that $S = I$ for $\omega = 0$, the solution of this differential equation for the matrix S is:

$$S = e^{\frac{\omega}{2}\gamma^1\gamma^4} = \cos \frac{\omega}{2} + \gamma_1\gamma_4 \sin \frac{\omega}{2}. \quad (19.22)$$

The last transformation depends on the fact that

$$(\gamma^1\gamma^4)^2 = -I,$$

for, in this case, the relation

$$e^{i\frac{\omega}{2}} = \cos \frac{\omega}{2} + i \sin \frac{\omega}{2}$$

will remain true if i is replaced by the matrix $\gamma_1\gamma_4$. From the result (19.22), we see that for a complete rotation ($\omega = 2\pi$), the matrix S does not return to its initial value, the unit matrix, but becomes $-(I)$:

$$S = -(I) \quad \text{for } \omega = 2\pi. \quad (19.22')$$

We are, therefore, dealing here with a *two-valued* representation of the rotation group of three-dimensional space, as we already know from the non-relativistic theory of spin. We shall come back to the connection between the matrices $T^{ik}(l, k = 1, 2, 3)$ and the angular momentum operators later.

We obtain in a similar way the transformation law for the ψ , i.e. the matrix S under the special Lorentz transformations which correspond to a relative motion of the co-ordinate systems in the x_1 -direction, and therefore, to rotations in the (x_1, x_4) -plane. We have now to replace x_2, γ^2 by x_4, γ^4 . If we consider the real time co-ordinate $x_0 = ct$ instead of $x_4 = ix_0$, we have to replace the angle of rotation $\omega = i\chi$ by real χ . Then

$$\begin{aligned} x'_1 &= x_1 \operatorname{Cosh} \chi - x_0 \operatorname{Sinh} \chi, & \operatorname{Tanh} \chi &= v/c \\ x'_0 &= -x_1 \operatorname{Sinh} \chi + x_0 \operatorname{Cosh} \chi, & \operatorname{Cosh} \chi &= \frac{1}{\sqrt{1-v^2/c^2}}, \quad \operatorname{Sinh} \chi = \frac{v/c}{\sqrt{1-v^2/c^2}}, \\ dx'_1 &= -d\chi x'_0, & dx'_0 &= -d\chi x'_1. \\ S &= e^{\frac{\chi}{2}i\gamma^1\gamma^4} = \operatorname{Cosh} \frac{\chi}{2} - i\gamma^1\gamma^4 \operatorname{Sinh} \frac{\chi}{2}. \end{aligned}$$

Summarising, we again stress that the existence of a four-current with the correct relativistic invariance properties on the one hand and a positive definite density on the other is an important feature of the Dirac theory. In an attempt to rewrite the Dirac wave equation in vector- or tensor form this property of the theory would be lost.

In the investigation of the behaviour of the wave function ψ under Lorentz transformations, no special representation of the matrices γ^k has been introduced. Different representations will be useful, according to the nature of the problem considered. In order to establish a connection with the mathematical literature, we choose now a representation in which γ^5 is made diagonal. [It is different from the representation given in eq. (18.19), in which $\beta = \gamma^4$ was chosen diagonal.] In the partitioned form in which the elements are themselves to be considered 2×2 matrices, we can set

$$\gamma^k = \begin{pmatrix} 0, & i\sigma_k \\ -i\sigma_k, & 0 \end{pmatrix} \quad \text{for } k = 1, 2, 3; \quad \gamma^4 = \begin{pmatrix} 0, & I \\ I, & 0 \end{pmatrix}; \quad \gamma^5 = \begin{pmatrix} -I, & 0 \\ 0, & I \end{pmatrix}. \quad (19.23)$$

[$I = 2 \times 2$ unit matrix and σ_k is defined by (18.17).] We now observe that according to eq. (19.20) the matrix S which effects the transformation

$$\psi' = S\psi$$

commutes with γ^4 , if the corresponding co-ordinate transformation has determinant +1. From this follows immediately that for these transformations which are called proper transformations, S assumes the form

$$S = \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma' \end{pmatrix},$$

where Σ and Σ' are 2×2 matrices. From eqs. (19.13) then follows further, with the value of γ^4 given in (19.23), that $\Sigma' = \tilde{\Sigma}^{-1}$. Therefore,

$$S = \begin{pmatrix} \Sigma & 0 \\ 0 & \tilde{\Sigma}^{-1} \end{pmatrix}. \quad (19.24)$$

The term T^{11} which determines the infinitesimal rotations will be, according to eqs. (19.16), (18.18) and (19.23), given by

$$T^{11} = \frac{1}{2} i \begin{pmatrix} \sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}; \dots \quad (19.25a)$$

$$T^{12} = \frac{1}{2} i \begin{pmatrix} \sigma_1 & 0 \\ 0 & -\sigma_1 \end{pmatrix}; \dots \quad (19.25b)$$

where the remaining matrices of the set T^{ij} can be obtained by cyclic permutation of the indices 1, 2, 3.

From this, it is seen that the four components ψ_1, \dots, ψ_4 break up into two pairs $(\psi_1, \psi_2; \psi_3, \psi_4)$, which under Lorentz transformations transform among themselves only and in fact the second pair transforms contragradiently with respect to the complex conjugates of the first pair, i.e.

$$\psi_1^* \psi_3 + \psi_2^* \psi_4$$

is an invariant. Our four-by-four representation of the proper (determinant +1) Lorentz group breaks up into two two-by-two representations. Quantities with two components φ_1, φ_2 which transform, under Lorentz transformations according to

$$\varphi' = \Sigma \varphi \quad (19.26)$$

are called spinors or also semi-vectors. Since the σ_k have the trace zero, the Σ always have the determinant 1. To the sub-group of the three-dimensional rotations corresponds a *unitary* matrix Σ (as was the case in the non-relativistic theory of spin also), whereas for the Lorentz transformations which do not leave the fourth co-ordinate x_4 unchanged, Σ is not unitary because of its imaginary character. By counting the parameters it is shown easily that the most general 2×2 matrix Σ with the determinant +1 belongs indeed to a particular Lorentz transformation.

The decomposition of the four-by-four representation of the Lorentz group into two-by-two representations fails if we consider the reflections also. In particular under the reflection considered earlier, eq. (19.21),

$$x'_k = -x_k \quad \text{for } k = 1, 2, 3; \quad x'_4 = x_4,$$

where, according to (19.21')

$$\psi' = \gamma^4 \psi$$

we obtain, on using the matrix for γ^4 given in (19.23)

$$\psi'_1 = \psi_2, \quad \psi'_2 = \psi_4, \quad \psi'_3 = \psi_1, \quad \psi'_4 = \psi_3,$$

i.e. the pairs (ψ_1, ψ_2) and (ψ_3, ψ_4) will be simply interchanged.

Calculations with quantities which behave under Lorentz transformations like ψ_1, ψ_2 , or ψ_3, ψ_4 have been extended by van der Waerden⁴ to a systematic "spinor calculus" which represents an extension of the ordinary tensor calculus and which enables all possible irreducible representations of the Lorentz group to be evaluated. We might observe here that this calculus is not always advantageous in spite of its formal completeness since the splitting up of all quantities with four components into two two-component ones, effected by specialisation to a γ^5 in the diagonal form, often introduces an unnecessary

⁴ B.L. van der Waerden, Göttinger Nachr. (1929), 100. For further applications, see G.E. Uhlenbeck and O. Laporte, Phys. Rev. 37, 1380 (1931).

complication of the formalism. For certain problems, other representations of the γ^k than the one given in (19.23) prove to be more useful, e.g. the one given in (18.19) where γ^k is diagonal.

We may now briefly mention the possibility of setting up Lorentz-invariant equations which contain only the two-component quantities ψ , introduced above.⁵ This rests on the fact that all covariant quantities constructed from matrices which commute with λ^i remain covariant, if we write down these quantities for a single two-component pair. We now return to the matrices defined by (19.1')

$$\alpha^k = i\gamma^4\gamma^k, \quad \beta = \gamma^4,$$

which are given, in our case, according to (19.23) by

$$\alpha^k = \begin{pmatrix} \sigma_k & 0 \\ 0 & -\sigma_k \end{pmatrix}, \quad \beta = \gamma^4 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}.$$

and to the original form

$$\frac{1}{c} \frac{\partial \psi}{\partial t} + \sum_{k=1}^3 \alpha^k \frac{\partial \psi}{\partial x_k} + i \frac{mc}{\hbar} \beta \psi = 0$$

of the Dirac equation, the current being given by

$$s_k = (\psi^* \alpha^k \psi); \quad s_4 = i(\psi^* \psi).$$

We see immediately that the quantities s_k , constructed from the two-component ψ according to the relations

$$s_k = (\psi^* \sigma_k \psi); \quad s_4 = i s_0 = i(\psi^* \psi) \quad (19.27)$$

form the components of a four-vector. These further satisfy the identity

$$s_0^2 = \sum_{k=1}^3 s_k^2. \quad (19.27')$$

We then see that the matrix β which does not commute with γ^k and which interchanges the pairs (ψ_1, ψ_2) and (ψ_3, ψ_4) cannot appear in a covariant wave equation built up from only two components ψ . The term in the wave equation containing the rest mass must, therefore, be absent in a two-component equation. Striking off this term signifies, physically, the transition to particles with rest-mass zero which always travel with the velocity of light, c and the energy and momentum of which are related by

$$\frac{E^2}{c^2} = \sum_k p_k^2 \quad (19.28)$$

as for photons. The two-component equation then reads

$$\frac{1}{c} \frac{\partial \psi}{\partial t} + \sum_{k=1}^3 \sigma_k \frac{\partial \psi}{\partial x_k} = 0. \quad (19.29)$$

It is relativistically invariant and leads, with the vector s_k defined by (19.27), to the continuity equation:

$$\sum_{k=1}^3 \frac{\partial s_k}{\partial x_k} = 0.$$

This wave equation is not invariant under reflections (i.e. interchange of left and right) as indeed follows from its derivation.⁶ The lack of invariance of the wave equation under reflections is exhibited in a characteristic coupling between the directions of spin and of the current, but we shall not go into this any further here. It may, however, be mentioned that eq. (19.29) possesses eigensolutions belonging to states of positive as well as of negative energies. For given values of energy and momentum, which satisfy (19.28), we, however, obtain here only one eigensolution.

⁵ This possibility was pointed out by H. Weyl, Z. Physik 56, 330 (1929).

⁶ This equation has recently been used for the neutrino to represent the non-conservation of parity in weak interactions.*

*See T.D. Lee and C.N. Yang, Phys. Rev. 105, 1671 (1957); L. Landau, Nucl. Phys. 3, 127 (1957); A. Salam, Nuovo Cimento 5, 299 (1957).

20. The Behaviour of Wave-Packets in the Free-Particle Case

Just as we obtained eqs. (3.1'') and (3.4), we obtain, by Fourier-decomposition, the momentum space eigenfunctions $\varphi_e(\vec{p}, t)$ from the co-ordinate space eigenfunctions $\psi_e(\vec{x}, t)$ according to

$$\varphi_e(\vec{p}, t) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int \psi_e(\vec{x}, t) e^{-\frac{i}{\hbar}(\vec{p}\cdot\vec{x})} d^3x, \quad (20.1)$$

$$\psi_e(\vec{x}, t) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int \varphi_e(\vec{p}, t) e^{+\frac{i}{\hbar}(\vec{p}\cdot\vec{x})} d^3p, \quad (20.1a)$$

where

$$W(p_1, p_2, p_3) dp_1 dp_2 dp_3 = \sum_k |\varphi_e(\vec{p}, t)|^2 dp_1 dp_2 dp_3 \quad (20.2)$$

is to be interpreted as the probability that the momentum components of the particle lie between p_k and $p_k + dp_k$. In the non-relativistic quantum mechanics the energy is uniquely determined by the momentum, whereas in relativistic mechanics, according to (18.1) both the energy values

$$\frac{1}{c} E = \pm \sqrt{m^2 c^2 + \sum_k p_k^2} \quad (20.3)$$

distinguished by their signs, are possible. From eq. (18.13) follows immediately the equation (for $\varphi_e(\vec{p})$)

$$-\frac{\hbar}{i} \frac{1}{c} \frac{\partial \varphi}{\partial t} = \left(\sum_{k=1}^3 \alpha^k p_k + mc\beta \right) \varphi, \quad (20.4)$$

where we use the matrix notation. The general solution of this equation reads

$$\left. \begin{aligned} \varphi_e &= C_e^{(+)} e^{-\frac{i}{\hbar} \sqrt{m^2 c^2 + \sum_k p_k^2} \cdot ct} \\ &\quad + C_e^{(-)} e^{+\frac{i}{\hbar} \sqrt{m^2 c^2 + \sum_k p_k^2} \cdot ct} \end{aligned} \right\} \quad (20.5)$$

where $C_e^{(+)}$ and $C_e^{(-)}$ satisfy the equations

$$+ \sqrt{m^2 c^2 + \sum_{k=1}^3 p_k^2} C^{(+)} = \left(\sum_{k=1}^3 \alpha^k p_k + mc\beta \right) C^{(+)}, \quad (20.6a)$$

$$- \sqrt{m^2 c^2 + \sum_{k=1}^3 p_k^2} C^{(-)} = \left(\sum_{k=1}^3 \alpha^k p_k + mc\beta \right) C^{(-)}. \quad (20.6b)$$

Each of these equations has again two linearly independent solutions. On account of the Hermiticity of α^k , we also have

$$+ \sqrt{m^2 c^2 + \sum_{k=1}^3 p_k^2} C^{*(+)}) = C^{*(+)} \left(\sum_{k=1}^3 \alpha^k p_k + mc\beta \right), \quad (20.6a^*)$$

$$- \sqrt{m^2 c^2 + \sum_{k=1}^3 p_k^2} C^{*(-)} = C^{*(-)} \left(\sum_{k=1}^3 \alpha^k p_k + mc\beta \right). \quad (20.6b^*)$$

Multiplication of eq. (20.6a) on the left by C^{*-1} and of eq. (20.6b*) on the right by

$C^{(+)}_k$ leads to the orthogonality condition

$$\sum_k C^{*(-)}_k C^{(+)}_k = \sum_k C^{(-)}_k C^{*(+)}_k = 0. \quad (20.7)$$

The probability $W(p_1, p_2, p_3)$ is, therefore, constant in time:

$$W(p_1, p_2, p_3) = \sum_k (|C^{(+)}_k|^2 + |C^{(-)}_k|^2) = \text{const.} \quad (20.8)$$

It may be noted that the doubts mentioned earlier against the existence of a probability density $W(x_1, x_2, x_3)$ in co-ordinate space affect the corresponding probability density in momentum space only to a small extent. The only problem is whether the momentum of a particle can be measured exactly in an arbitrarily short interval of time [cf. inequality (2.11)]; on the contrary, it is clear that the momentum can be measured with any desired degree of accuracy in a sufficiently long time. For a free-particle wave-packet, for which the momenta are constant in time, $W(p_1, p_2, p_3)$ can, therefore, be determined exactly. But we can say something more also: Even if a free particle is subject to a force (interaction with other particles or with radiation) only during a finite interval of time, the momentum of the particle can be measured with arbitrary accuracy both before and after the interaction. The velocity distribution of particles after collision is, therefore, an exact and rigorously meaningful concept also in relativistic quantum mechanics. So also is the intensity distribution of the radiation scattered by a particle as a function of the frequency and direction, as we shall see later.

After this digression, let us investigate the consequences of the existence of the solutions $\psi_k^{(-)}$ which belong to the negative energy, for the behaviour of the wave-packet. To start with it follows, in contrast to the non-relativistic theory, that the total current J_k which according to eqs. (18.16') and (20.1) is given by

$$\frac{1}{c} J_k = \int (\psi^* \alpha^k \psi) d^3 x = \int (\psi^* \alpha^k \varphi) d^3 p \quad (20.9)$$

is no longer constant in time. Rather, it is given according to (20.5), by

$$\left. \begin{aligned} \frac{1}{c} J_k &= \int (C^{(+)} \alpha^k C^{(+)}) d^3 p + \int (C^{(-)} \alpha^k C^{(-)}) d^3 p + \\ &+ \int [(C^{(+)} \alpha^k C^{(-)}) e^{\frac{2i}{\hbar} \sqrt{m^2 c^2 + \sum p_i^2} \cdot ct} + (C^{(-)} \alpha^k C^{(+)}) e^{-\frac{2i}{\hbar} \sqrt{m^2 c^2 + \sum p_i^2} \cdot ct}] d^3 p. \end{aligned} \right\} \quad (20.10)$$

The first two terms in eq. (20.10) are constant in time and correspond to what one should expect in analogy with classical relativistic mechanics. We determine their magnitude by multiplying (20.6a*) by $\alpha^l C^{(+)}$ from the right and (20.6b) by $C^{*-1} \alpha^l$ from the left. By addition, we obtain, remembering the anticommutation relations (I),

$$\begin{aligned} \sqrt{m^2 c^2 + \sum p_i^2} (C^{(+)} \alpha^l C^{(+)}) &= p_l (C^{(+)} C^{(+)}) , \\ -\sqrt{m^2 c^2 + \sum p_i^2} (C^{(-)} \alpha^l C^{(-)}) &= p_l (C^{(-)} C^{(-)}). \end{aligned}$$

On introducing the energy

$$E = \pm c \sqrt{m^2 c^2 + \sum p_i^2}$$

and the particle velocity

$$v_k = \frac{\partial E}{\partial p_k} = \frac{c^2 p_k}{E} = \frac{\pm c p_k}{\sqrt{m^2 c^4 + \sum p_k^2}} \quad (20.11)$$

which coincides with the group velocity of the waves, therefore, follows for the constant part of J_k the expression

$$\bar{J}_k = \int v_k^{(+)}(p) (C^{*(+)}) C^{(+)} d^3 p + \int v_k^{(-)}(p) (C^{*(-)}) C^{(-)} d^3 p. \quad (20.12)$$

The oscillation with the frequency $2|E|/\hbar$ which is superimposed on this* was called by Schrödinger "Zitterbewegung" (Jittery motion). It is a consequence of the interference between the positive and negative energy parts of the wave-packet; in wave-packets, which contain the eigenfunctions belonging to only one sign of the energy, this Zitterbewegung drops out.

Just as in the current, Zitterbewegung also shows up in the motion of the wave-packet whose mid-point is defined through

$$\bar{x}_k = \int x_k W(x) d^3 x = -\frac{\hbar}{i} \int \sum_q \varphi_q^* \frac{\partial \varphi_q}{\partial p_k} d^3 p. \quad (20.13)$$

This is so because, from the continuity equation (18.15), it follows directly that

$$\frac{d \bar{x}_k}{dt} = J_k. \quad (20.14)$$

Therefore, corresponding to the constant part of J_k , there is a uniform motion of \bar{x}_k and corresponding to the oscillating part, we have an oscillation of x_k . At first we might think of introducing as a subsidiary condition of the theory, that only such wave-packets are to be allowed which contain eigenfunctions belonging exclusively to positive energies. This is in fact possible if we have only the free-particle case in view, but when forces are present, it cannot be carried out in agreement with relativistic invariance and in correspondence with the classical-relativistic mechanics (cf. Sec. 24).

The mathematical formulation of the behaviour of general wave-packets and their properties with the passage of time becomes clearer if we change over from wave functions to operators.¹ This transition takes place exactly as in non-relativistic quantum mechanics; the only difference is that the operators act now also on the index p which can assume four values. In addition to an integration over the position or momentum co-ordinates, we have also to sum over p . If D is an operator which acts on the function $u_p(x)$, then

$$(D u)_q = \sum_{\sigma} D_{q\sigma} u_{\sigma}.$$

The operator is Hermitian, if for arbitrary functions u_p and v_p , we have

$$\int \sum_q (D u)_q^* u_q d^3 x = \int \sum_q v_q^* (D u)_q d^3 x. \quad (20.15)$$

*viz., the third and fourth terms of eq. (20.10).

¹ E. Schrödinger, Berl. Ber. (1930), 418; (1931), 63. — V. Fock, Z. Physik 55, 127 (1929); 68, 527 (1931). (In this paper are also to be found applications to the case when forces are present.) See, further, the discussion in E. Schrödinger, Z. Physik 70, 808 (1931); V. Fock, Z. Physik 70, 811 (1931).

The change of the operators with time is defined by the requirement that for two arbitrary solutions $\psi_p(t)$, $\psi'_p(t)$ of the wave equation, we should have

$$\int \sum_p \psi_p^*(t) [\mathbf{D}(0) \psi'(t)]_p d^3x = \int \sum_p \psi_p^*(0) [\mathbf{D}(t) \psi'(0)]_p d^3x.$$

Then

$$\mathbf{D}(t) = e^{\frac{i}{\hbar} \mathbf{H} t} \mathbf{D}(0) e^{-\frac{i}{\hbar} \mathbf{H} t}.$$

Therefore

$$\frac{d\mathbf{D}}{dt} = \frac{i}{\hbar} (\mathbf{H}\mathbf{D} - \mathbf{D}\mathbf{H}). \quad (20.16)$$

In fact for

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \mathbf{H}\psi; \quad -\frac{\hbar}{i} \frac{\partial \psi'}{\partial t} = \mathbf{H}\psi'.$$

where \mathbf{H} denotes the Hamiltonian, we have, remembering (20.15) and in case \mathbf{H} is Hermitian,

$$\frac{d}{dt} \int \sum_p \psi_p^* (\mathbf{D}\psi)_p d^3x = \frac{i}{\hbar} \int \sum_p \psi_p^* [(\mathbf{H}\mathbf{D} - \mathbf{D}\mathbf{H})\psi']_p d^3x. \quad (20.16')$$

Here it is assumed that \mathbf{D} does not contain the time explicitly.

The Hamiltonian of the free-particle Dirac equation is now given by

$$\mathbf{H} = c \left(\sum_{k=1}^3 \alpha^k \mathbf{p}_k + \beta m c \right),$$

where the α^k , β commute with the \mathbf{p}_k and \mathbf{x}_k and satisfy (I) and where we have further

$$\mathbf{p}_i \mathbf{x}_k - \mathbf{x}_k \mathbf{p}_i = \frac{\hbar}{i} \delta_{ik} I. \quad (20.17)$$

Now we find

$$\mathbf{x}_k = \frac{i}{\hbar} (\mathbf{H} \mathbf{x}_k - \mathbf{x}_k \mathbf{H}) = c \alpha_k, \quad (20.18a)$$

$$\mathbf{p}_k = \frac{i}{\hbar} (\mathbf{H} \mathbf{p}_k - \mathbf{p}_k \mathbf{H}) = 0, \quad (20.18b)$$

$$\alpha^k = \frac{i}{\hbar} (\mathbf{H} \alpha^k - \alpha^k \mathbf{H}) = \frac{2i}{\hbar} (c \mathbf{p}_k - \alpha^k \mathbf{H}) = \frac{2i}{\hbar} (\mathbf{H} \alpha^k - c \mathbf{p}_k), \quad (20.18c)$$

$$\beta = \frac{i}{\hbar} (\mathbf{H} \beta - \beta \mathbf{H}) = \frac{2i}{\hbar} (m c^2 - \beta \mathbf{H}) = \frac{2i}{\hbar} (\mathbf{H} \beta - m c^2), \quad (20.18d)$$

The last two relations follow since

$$\mathbf{H} \alpha^k + \alpha^k \mathbf{H} = 2c \mathbf{p}_k, \quad \mathbf{H} \beta + \beta \mathbf{H} = 2m c^2. \quad (20.19)$$

Specially noteworthy is the eq. (20.18a) which shows that $c\alpha^k$ play formally the role of velocities and that these are not directly connected with the momenta as in classical mechanics. This fact was first pointed out by Breit.² Eq. (20.18c) determines, according to (20.16'), the change with time of the total current.

The removal of the connection between velocity and momentum is closely related to the possibility of states of negative energy. To see this, let us, with Schrödinger, first decompose the wave-packet into positive and negative parts. We mean by positive function, the part built exclusively from $C_p^{(+)}$:

$$\psi_p^{(+)} = \int C_p^{(+)}(\vec{p}) e^{\frac{i}{\hbar} \vec{p} \cdot \vec{x}} e^{-\frac{i}{\hbar} \sqrt{m^2 c^2 + \sum p_k^2} ct} d^3p. \quad (20.20a)$$

Similarly,

$$\psi_p^{(-)} = \int C_p^{(-)}(\vec{p}) e^{\frac{i}{\hbar} \vec{p} \cdot \vec{x}} e^{-\frac{i}{\hbar} \sqrt{m^2 c^2 + \sum p_k^2} ct} d^3p. \quad (20.20b)$$

² G. Breit, Proc. Nat. Acad. Sci. U.S.A. 14, 553 (1928); cf. also Proc. Nat. Acad. Sci. U.S.A. 17, 70 (1931).

Here $C_p^{(+)}$ and $C_p^{(-)}$ satisfy the eqs. (20.6a) and (20.6b) respectively. Now we introduce the operation Λ , defined by

$$\Lambda = \frac{\sum_k \alpha^k p_k + \beta m c}{\sqrt{m^2 c^2 + \sum_k p_k^2}}. \quad (20.21)$$

The operator Λ is defined in momentum space and its action on the function $\psi_p(p)$ is clear. Λ is Hermitian and its square is 1:

$$\Lambda^2 = 1. \quad (20.22)$$

Therefore, Λ is unitary and commutes with the p_k and with the Hamiltonian. Evidently Λ takes $\psi_p^{(+)}$ into itself and each $\psi_p^{(-)}$ into its negative:

$$\Lambda \psi_p^{(+)} = \psi_p^{(+)}; \quad \Lambda \psi_p^{(-)} = -\psi_p^{(-)}. \quad (20.23)$$

Therefore,

$$\Lambda \varphi_p = \varphi_p^{(+)} - \varphi_p^{(-)}, \quad \varphi_p = \varphi_p^{(+)} + \varphi_p^{(-)}. \quad (20.24)$$

It is not difficult now to define Λ in co-ordinate space also. The operator p_k is simply $\frac{\hbar}{i} \frac{\partial}{\partial x_k}$. We have only to define the operator $\frac{1}{\sqrt{m^2 c^2 + \sum_k p_k^2}}$. For the function

$$\psi_q(x) = e^{\frac{i}{\hbar} \hat{p} \cdot \vec{x}}$$

this operator is already defined. Therefore,

$$\frac{1}{\sqrt{m^2 c^2 + \sum_k p_k^2}} \int \Lambda_q(p) e^{\frac{i}{\hbar} \hat{p} \cdot \vec{x}} d^3 p = \int \frac{\Lambda_q(p)}{\sqrt{m^2 c^2 + \sum_k p_k^2}} e^{\frac{i}{\hbar} \hat{p} \cdot \vec{x}} d^3 p.$$

Since

$$\Lambda_q(p) = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}} \int \psi_q(x') e^{-\frac{i}{\hbar} \hat{p} \cdot \vec{x}'} d^3 x',$$

on introducing the function

$$D(x) = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}} \iiint \frac{e^{-\frac{i}{\hbar} \hat{p} \cdot \vec{x}}}{\sqrt{m^2 c^2 + \sum_k p_k^2}} dp_1 dp_2 dp_3 = \frac{2\pi}{(2\pi\hbar)^{\frac{1}{2}}} \frac{\hbar}{i} \frac{1}{r} \int \frac{e^{-\frac{i}{\hbar} \hat{p} \cdot r} - e^{+\frac{i}{\hbar} \hat{p} \cdot r}}{\sqrt{m^2 c^2 + \hat{p}^2}} \hat{p} d\hat{p},$$

we have

$$\frac{1}{\sqrt{m^2 c^2 + \sum_k p_k^2}} \psi_q(\vec{x}) = \int D(\vec{x} - \vec{x}') \psi_q(\vec{x}') d^3 x'.$$

We need not go more closely into the evaluation of the function $D(x)$ which is singular at $r = 0$. It must, however, be stressed that Λ has then the property

$$\Lambda \psi_q^{(+)} = \psi_q^{(+)}, \quad \Lambda \psi_q^{(-)} = -\psi_q^{(-)}, \quad (20.23')$$

$$\psi_q^{(+)} = \frac{1}{2}(1 + \Lambda) \psi_q, \quad \psi_q^{(-)} = \frac{1}{2}(1 - \Lambda) \psi_q$$

in co-ordinate space also. We observe, further, that

$$\frac{1}{\sqrt{m^2 c^2 + \sum_k p_k^2}} \Lambda = \frac{\sum_k \alpha^k p_k + \beta m c}{m^2 c^2 + \sum_k p_k^2} = c H^{-1},$$

since this gives the identity operator when multiplied by $\frac{1}{c} H$.

We can now split every operator D into an even and an odd part by means of this operator Λ . Here an even operator is one which changes every positive (or negative) function again into a positive (or negative) function; an odd operator is one which changes every positive (or negative) function into a negative (or positive) function. Since all the positive functions are orthogonal to all the negative func-

tions, the odd operators have the expectation value zero in all states which can be represented by wave-packets with only positive or negative energies. Now

$$\begin{aligned} \mathbf{A} \mathbf{D} \mathbf{A} \psi^{(+)} &= \mathbf{A} \mathbf{D} \psi^{(+)} = (\mathbf{D} \psi^{(+)})^{(+)} - (\mathbf{D} \psi^{(+)})^{(-)}, \\ \mathbf{A} \mathbf{D} \mathbf{A} \psi^{(-)} &= -\mathbf{A} \mathbf{D} \psi^{(-)} = -(\mathbf{D} \psi^{(-)})^{(+)} + (\mathbf{D} \psi^{(-)})^{(-)}. \end{aligned}$$

Therefore,

$$\frac{1}{2}(\mathbf{D} + \mathbf{A} \mathbf{D} \mathbf{A}) = \frac{1}{2} \mathbf{A} (\mathbf{A} \mathbf{D} + \mathbf{D} \mathbf{A}) = g(\mathbf{D}) \quad (20.25 \text{ a})$$

is the even part and

$$\frac{1}{2}(\mathbf{D} - \mathbf{A} \mathbf{D} \mathbf{A}) = \frac{1}{2}(\mathbf{D} \mathbf{A} - \mathbf{A} \mathbf{D}) \mathbf{A} = u(\mathbf{D}) \quad (20.25 \text{ b})$$

is the odd part of \mathbf{D} . Under the unitary transformation

$$\mathbf{D} \rightarrow \mathbf{A} \mathbf{D} \mathbf{A}$$

$\mathbf{D} = g + u$ transforms into $g - u$.

It is now easy to evaluate the even part of α^k and β . We find

$$g(\alpha^k) = c \mathbf{H}^{-1} \mathbf{p}_k, \quad g(\beta) = m c^2 \mathbf{H}^{-1}. \quad (20.26)$$

The even parts of α^k and β have, therefore, again the classical connection with momentum and energy which was not the case for the original operators α^k and β . Further, remembering that

$$\mathbf{x}_k \mathbf{A} - \mathbf{A} \mathbf{x}_k = -\frac{\hbar}{i} \frac{\partial \mathbf{A}}{\partial \mathbf{p}_k}$$

for the odd part of \mathbf{x}_k , we find:

$$u(\mathbf{x}_k) = \frac{\hbar}{2i} c \mathbf{H}^{-1} (\alpha^k - \mathbf{p}_k c \mathbf{H}^{-1}) = \frac{\hbar}{2i} c \mathbf{H}^{-1} u(\alpha^k) \quad (20.27)$$

in agreement with the expressions given by Schrödinger.

21. The Wave Equation when Forces are Present

In classical relativistic mechanics of a particle, the Hamiltonian of a charged particle under the influence of an external electromagnetic field is obtained by replacing the energy E by $E + e\Phi_0$ and the momenta p_k by $p_k + \frac{e}{c}\Phi_k$. Here the charge of the particle is taken to be $-e$, which is the case for an electron, Φ_0 is the (electric) scalar and Φ_k the (magnetic) vector potential of the external field. If we combine $(\Phi_k, i\Phi_0)$ into the four-vector potential Φ , and $(p_k, i\frac{E}{c})$ into the four-vector p , of momentum and energy, the two sets of transformations above can be formulated as replacing p_v by $p_v + \frac{e}{c}\Phi_v$. Dirac retains this ansatz in quantum theory and sets, therefore, for the wave equation of a charged particle under the influence of external forces, as a generalisation of (II) of Section 19,

$$\sum_{\mu=1}^4 \gamma^\mu \left(\frac{\hbar}{i} \frac{\partial}{\partial x_\mu} + \frac{e}{c} \Phi_\mu \right) \psi - imc\psi = 0 \quad (\text{III})$$

or

$$\sum_{\mu} \gamma^\mu \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} \Phi_\mu \right) \psi + \frac{mc}{\hbar} \psi = 0. \quad (\text{III}')$$

For the functions ψ^\dagger defined by (19.2), we then have

$$\sum_{\mu} \left(\frac{\partial}{\partial x_\mu} - \frac{ie}{\hbar c} \Phi_\mu \right) \psi^\dagger \gamma^\mu - \frac{mc}{\hbar} \psi^\dagger = 0. \quad (\text{III}^\dagger)$$

The definitions (19.4), (18.11) and (18.16) of the four-vectors can, therefore, be retained as such since the continuity equation

$$\sum_{\nu=1}^4 \frac{\partial s_\nu}{\partial x_\nu} = 0$$

is a consequence of the wave equation. Further, the considerations regarding the relativistic invariance are not changed.

An important factor is that only the field strengths

$$F_{\mu\nu} = \frac{\partial \Phi_\nu}{\partial x_\mu} - \frac{\partial \Phi_\mu}{\partial x_\nu} \quad (21.1)$$

have a direct physical significance; the potentials, therefore, are defined only up to an additive gradient of an arbitrary function $f(x_1 \dots x_4)$. If Φ_μ is replaced by

$$\Phi'_\mu = \Phi_\mu + \frac{\partial f}{\partial x_\mu}, \quad (21.2a)$$

$F_{\mu\nu}$ remains unchanged. This substitution, introduced in (III) above shows that in the transition from Φ_μ to Φ'_μ , the wave functions ψ_ν themselves transform according to

$$\psi'_\nu = \psi_\nu e^{-\frac{if}{\hbar c}}, \quad (21.2b)$$

For, then

$$\left(\frac{\partial}{\partial x_\mu} + \frac{i\epsilon}{\hbar c} \Phi'_\mu \right) \psi' = \left\{ \left(\frac{\partial}{\partial x_\mu} + \frac{i\epsilon}{\hbar c} \Phi_\mu \right) \psi \right\} e^{-\frac{if}{\hbar c}}. \quad (21.3)$$

The transformations (21.2a) and (21.2b) are called *gauge transformations* because of a similar situation in an earlier theory of gravitation and electricity due to Weyl. The invariance of the wave equation under these transformations is called *gauge invariance*.¹

Going over from γ^ν to the α^k and β , the wave equation can also be written as

$$\frac{1}{c} \frac{\partial \psi}{\partial t} - \frac{i\epsilon}{\hbar c} \Phi_0 \psi + \sum_{k=1}^3 \alpha^k \left(\frac{\partial \psi}{\partial x_k} + \frac{i\epsilon}{\hbar c} \Phi_k \psi \right) + i \frac{mc}{\hbar} \beta \psi = 0, \quad (21.4)$$

so that the Hamiltonian, defined by

$$\frac{\hbar}{i} \frac{\partial \psi}{\partial t} + H\psi = 0$$

is given by

$$H = -e\Phi_0 + c \left[\sum_{k=1}^3 \alpha^k \left(p_k + \frac{e}{c} \Phi_k \right) + mc\beta \right] \quad (21.5)$$

If we introduce, further, instead of p_k and H the gauge-invariant operators

$$\left. \begin{aligned} \pi_k &= p_k + \frac{e}{c} \Phi_k = \frac{\hbar}{i} \frac{\partial}{\partial x_k} + \frac{e}{c} \Phi_k, \\ \pi_0 &= \frac{H}{c} + \frac{e}{c} \Phi_0 = -\frac{\hbar}{i} \frac{1}{c} \frac{\partial}{\partial t} + \frac{e}{c} \Phi_0, \\ \pi_4 &= i\pi_0 = \frac{\hbar}{i} \frac{\partial}{\partial x_4} + \frac{e}{c} \Phi_4 \end{aligned} \right\} \quad (21.6)$$

¹ F. London, Z. Physik 42, 375 (1927). – H. Weyl, Z. Physik 56, 330 (1929).

we have

$$\sum_{\mu=1}^4 \gamma^\mu \pi_\mu \psi - i m c \psi = 0. \quad (\text{III}'')$$

The operators π_μ satisfy the commutation relations

$$\pi_\mu \pi_\nu - \pi_\nu \pi_\mu = \frac{\hbar}{i} \frac{e}{c} F_{\mu\nu}, \quad (21.7)$$

whereas the operators p_μ commute among themselves.

The requirements of relativistic invariance, gauge invariance and the correspondence principle do not, however, determine the ansatz (III) quite uniquely. There remains the possibility of modifying it by adding the term

$$\frac{1}{c} M \sum_{\mu} \sum_{\nu} F_{\mu\nu} \gamma^\mu \gamma^\nu \psi$$

where $F_{\mu\nu}$ denote again the field strengths of the external field and M is a real constant with the dimension of charge times length. It is true, however, that even without such an additional term the wave equation of the first order is in agreement with experiments. *In fact this equation by itself gives the spin of the electron including the absolute value $eh/2mc$ of its magnetic moment.* This is to be considered as an important success of the Dirac equation. The above statement is readily verified if we go over from (III) to the wave equation of the second order by means of the same operation which led to the eq. (18.5) in the free-particle case. Thus applying the operator

$$\sum_{\mu=1}^4 \gamma^\mu \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} \Phi_\mu \right) - \frac{mc}{\hbar}$$

from the left on (III'), we immediately get

$$\sum_{\mu} \sum_{\nu} \gamma^\mu \gamma^\nu \left[\left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} \Phi_\mu \right) \left(\frac{\partial}{\partial x_\nu} + \frac{ie}{\hbar c} \Phi_\nu \right) - \frac{m^2 c^2}{\hbar^2} \right] \psi = 0.$$

Here we separate the terms with $\mu = \nu$ from the terms with $\mu \neq \nu$. The former give, in agreement with earlier relativistic theories [since $(\gamma^\mu)^2 = 1$],

$$\sum_{\nu} \left(\frac{\partial}{\partial x_\nu} + \frac{ie}{\hbar c} \Phi_\nu \right)^2 \psi.$$

The latter give (since $\gamma^\mu \gamma^\nu = -\gamma^\nu \gamma^\mu$ for $\mu \neq \nu$)

$$\begin{aligned} \frac{1}{2} \sum_{\mu} \sum_{\nu} \gamma^\mu \gamma^\nu & \left\{ \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} \Phi_\mu \right) \left(\frac{\partial}{\partial x_\nu} + \frac{ie}{\hbar c} \Phi_\nu \right) - \left(\frac{\partial}{\partial x_\nu} + \frac{ie}{\hbar c} \Phi_\nu \right) \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} \Phi_\mu \right) \right\} \psi \\ & = \frac{1}{2} \sum_{\mu} \sum_{\nu} \gamma^\mu \gamma^\nu \frac{ie}{\hbar c} F_{\mu\nu} \psi, \end{aligned}$$

in which the field strengths are again given by eq. (21.1). The end result is, therefore,

$$\left[\sum_{\nu} \left(\frac{\partial}{\partial x_\nu} + \frac{ie}{\hbar c} \Phi_\nu \right)^2 - \frac{m^2 c^2}{\hbar^2} \right] \psi + \frac{1}{2} \sum_{\mu} \sum_{\nu} \gamma^\mu \gamma^\nu \frac{ie}{\hbar c} F_{\mu\nu} \psi = 0 \quad (21.8)$$

or on introducing the operators $p_\nu = \frac{\hbar}{i} \frac{\partial}{\partial x_\nu}$:

$$\left[\sum_{\nu} \left(p_\nu + \frac{e}{c} \Phi_\nu \right)^2 + m^2 c^2 \right] \psi + \frac{1}{2} \sum_{\mu} \sum_{\nu} \gamma^\mu \gamma^\nu \frac{ie}{\hbar c} F_{\mu\nu} \psi = 0. \quad (21.8')$$

The fact that the last term of the above equation contains the interaction energy of the spin with the external field will become clear from the considerations of the following section where it will be shown in general that the non-relativistic quantum mechanics of spin, as developed in Sec. 13, is contained as an approximation, in Dirac's relativistic theory.

The treatment of the radiation processes based on the correspondence principle can also be carried over directly to the Dirac theory. Only, in eq. (15.28) for the perturbation energy of radiation, the expression

$$i_s = (-e) c (\psi^* \alpha^s \psi)$$

must be substituted for the current instead of the non-relativistic expression (4.18). The circumstance that in the Dirac theory, the four-potential is not present in the current and is present in the Hamiltonian only linearly proves a simplifying feature in many calculations as compared with the non-relativistic theory.

The most important applications of the Dirac theory which lead to consequences, which are characteristic of this theory and which can be tested experimentally are: (1) The computation of the exact eigenvalues of an electron in a Coulomb central field which turn out to be identical with the ones evaluated by Sommerfeld in his theory of the relativistic fine structure of atoms² and (2) the formula of Klein and Nishina³ for the intensity of the scattered radiation of short wave-length by free electrons.

Besides the law of conservation of charge, there is also a law of conservation of energy-momentum. This does not simply mean the conservation of the energy and momentum of a matter-field alone. This proves true only in the free-particle case whereas in the presence of an electromagnetic field, momentum and energy will be given to or drawn out of the system. However, there exists, in general, an energy-momentum tensor $T_{\mu\nu}$, which satisfies the relation

$$\sum_{\nu=1}^4 \frac{\partial T_{\mu\nu}}{\partial x_\nu} = \sum_\nu F_\mu s_\nu = (-e) i \sum_\nu F_{\mu\nu} (\psi^\dagger \gamma^\nu \psi). \quad (21.9)$$

By an elementary calculation it can in fact be easily verified that this relation is a consequence of the wave equation, if⁴

$$\left. \begin{aligned} -\frac{i}{c} T_{\mu\nu} &= \frac{1}{2} \left\{ \psi^\dagger \gamma^\nu \left[\left(p_\mu + \frac{e}{c} \Phi_\mu \right) \psi \right] - \left[\left(p_\mu - \frac{e}{c} \Phi_\mu \right) \psi^\dagger \right] \gamma^\nu \psi \right\} \\ &= \frac{1}{2} \frac{\hbar}{i} \left(\psi^\dagger \gamma^\nu \frac{\partial \psi}{\partial x_\mu} - \frac{\partial \psi^\dagger}{\partial x_\mu} \gamma^\nu \psi \right) + \frac{e}{c} \Phi_\mu (\psi^\dagger \gamma^\nu \psi) \end{aligned} \right\} \quad (21.10)$$

The second term of the last line of (21.10) gives the components $T_{\mu\nu}$ the correct reality properties. Here the commutation relations of the Dirac matrices have not yet been used. But we can utilise them to sym-

² W. Gordon, Z. Physik **48**, 11 (1928). – C.G. Darwin, Proc. Roy. Soc. Lond., Ser. A **118**, 654 (1928). For further details see H.A. Bethe and E.E. Salpeter, Quantum Mechanics of One- and Two-Electron Atoms, Springer-Verlag, Berlin, (1957).

³ O. Klein and Y. Nishina, Z. Physik **52**, 853 (1929). – Y. Nishina, Z. Physik **52**, 869 (1929); cf. also J. Waller, Z. Physik **58**, 75 (1929). Further details are to be found in G. Källén, Quantum Electrodynamics, loc. cit.

⁴ In the earlier relativistic quantum theory (cf. footnote 3, Sec. 18), the energy momentum tensor was introduced by E. Schrödinger, Ann. d. Phys. **82**, 265 (1927). The calculation for the Dirac theory, which is similar, is to be found in F. Möglich, Z. Physik **48**, 852 (1928). In his comprehensive treatment of the question, H. Tetrode [Z. Physik **49**, 858 (1928)] has shown the possibility of symmetrising the tensor.

metrise the energy-momentum tensor. To show this, we multiply the wave equation (III) on the left by $\psi^\dagger \gamma^\mu \gamma^\nu$ the wave equation (III \dagger) on the right by $\gamma^\mu \gamma^\nu \psi$ and add. We then obtain

$$\sum_{\rho=1}^4 \left\{ \psi^\dagger \gamma^\mu \gamma^\nu \gamma^\rho \left[\left(p_\rho + \frac{e}{c} \Phi_\rho \right) \psi \right] + \left[\left(p_\rho - \frac{e}{c} \Phi_\rho \right) \psi^\dagger \right] \gamma^\rho \gamma^\mu \gamma^\nu \psi \right\} = 0.$$

Only the case $\mu \neq \nu$ is of interest to us, which we shall from now on assume explicitly. Then we separate the terms with $\rho = \mu$ and $\rho = \nu$ on the one hand and the terms with $\rho \neq \mu, \rho \neq \nu$ on the other, the latter being characterised by a prime on the summation signs. Remembering the anticommutation relations (I') of the matrices γ^μ , we then obtain

$$\frac{2i}{c} (T_{\mu\nu} - T_{\nu\mu}) + \sum'_{\substack{\rho \neq \mu \\ \rho \neq \nu}} \left\{ \psi^\dagger \gamma^\mu \gamma^\nu \gamma^\rho \left[\left(p_\rho + \frac{e}{c} \Phi_\rho \right) \psi \right] + \left[\left(p_\rho - \frac{e}{c} \Phi_\rho \right) \psi^\dagger \right] \gamma^\mu \gamma^\nu \gamma^\rho \psi \right\} = 0$$

or

$$\frac{2}{c} (T_{\mu\nu} - T_{\nu\mu}) = \hbar \sum'_{\substack{\rho \neq \mu \\ \rho \neq \nu}} \frac{\partial}{\partial x_\rho} (\psi^\dagger \gamma^\mu \gamma^\nu \gamma^\rho \psi). \quad (21.11)$$

Since γ^ν and γ^ρ anticommute for $\rho \neq \nu$, the divergence of $T_{\mu\nu} - T_{\nu\mu}$ vanishes:

$$\sum_{\nu=1}^4 \frac{\partial}{\partial x_\nu} (T_{\mu\nu} - T_{\nu\mu}) = 0. \quad (21.11')$$

If we, therefore, form the symmetric tensor

$$\Theta_{\mu\nu} = \Theta_{\nu\mu} = \frac{1}{2} (T_{\mu\nu} + T_{\nu\mu}) = T_{\mu\nu} - \frac{\hbar c}{4} \sum'_{\substack{\rho \neq \mu \\ \rho \neq \nu \\ (\mu \neq \nu)}} \frac{\partial}{\partial x_\rho} (\psi^\dagger \gamma^\mu \gamma^\nu \gamma^\rho \psi) \quad (21.12)$$

(for $\mu = \nu$ the last term is to be dropped), then it always satisfies a relation

$$\sum_{\nu=1}^4 \frac{\partial \Theta_{\mu\nu}}{\partial x_\nu} = \sum_\nu F_{\mu\nu} s_\nu = (-e) i \sum_\nu F_{\mu\nu} (\psi^\dagger \gamma^\nu \psi), \quad (21.13)$$

which is of the form (21.9).

As special applications of the relations (21.9) and (21.13) we may mention the conservation laws of momentum and angular momentum. By integrating by parts, we obtain, for $k = 1, 2, 3$,

$$J_k = \frac{1}{ic} \int T_{k\ell} d^3 x = \frac{1}{ic} \int \Theta_{k\ell} d^3 x = \int (\psi^\dagger \gamma^\ell \pi_k \psi) d^3 x = \int (\psi^\ell \pi_k \psi) d^3 x \quad (21.14)$$

and

$$\dot{J}_k = (-e) \int \sum_\nu F_{k\nu} s_\nu d^3 x, \quad (21.15)$$

remembering (19.2). This is, according to eq. (20.16), equivalent to the operator-relation

$$\frac{d}{dt} \pi_k = \frac{e}{c} \frac{\partial \Phi_k}{\partial t} + \frac{i}{\hbar} (\mathbf{H} \pi_k - \pi_k \mathbf{H}) = (-e) \left(\mathcal{E}_k + \sum_{i=1}^3 F_{ki} \alpha^i \right), \quad (21.15')$$

which can be derived from (21.5) directly. Further from eq. (21.13), the law of conservation of angular momentum follows. With

$$P_{ik} = -P_{ki} = \frac{1}{ic} \int (x_i \Theta_{ki} - x_k \Theta_{ii}) d^3 x, \quad (i, k = 1, 2, 3), \quad (21.16)$$

and

$$d_{ik} = (-e) \sum_\nu (x_i F_{k\nu} - x_k F_{i\nu}) s_\nu, \quad (21.17)$$

or as operator relation

$$\mathbf{d}_{ik} = (-e) \left[x_i \mathcal{E}_k - x_k \mathcal{E}_i + \sum_{l=1}^3 (x_i F_{kl} - x_k F_{il}) \alpha^l \right] \quad (21.17')$$

we have

$$\dot{\mathbf{P}}_{ik} = \int d_{ik} d^3x. \quad (21.18)$$

In deriving the last relation from (21.13) the symmetry of the tensor Θ_{ik} has been essentially made use of.

By substituting from eqs. (21.10) and (21.12) in (21.16), the expression for the components of the angular momentum can be further transformed. We thus obtain

$$\mathbf{P}_{ik} = \frac{1}{i} \int \psi^\dagger \gamma^k (x_i \pi_k - x_k \pi_i) \psi d^3x + \frac{\hbar}{2i} \int (\psi^\dagger \gamma^k \gamma^i \gamma^k \psi) d^3x$$

or according to eq. (19.2) we have

$$\mathbf{P}_{ik} = \int \psi^\dagger \left[(x_i \pi_k - x_k \pi_i) + \frac{\hbar}{2i} \alpha^i \alpha^k \right] \psi d^3x. \quad (21.19)$$

The relation (21.18) can also then be written as the operator relation

$$\mathbf{P}_{ik} = x_i \pi_k - x_k \pi_i + \frac{\hbar}{2i} \alpha^i \alpha^k; \quad \frac{d}{dt} \mathbf{P}_{ik} = \mathbf{d}_{ik} \quad (21.20)$$

The last equation can also be derived directly by taking the commutator of the operator on the left with \mathbf{H} . The occurrence of the additional term $\frac{\hbar}{2i} \alpha^i \alpha^k$ in the angular momentum operator is closely connected with the behaviour of ψ under infinitesimal rotations, this being due to the nature of the operator $\gamma^i \gamma^k = i\alpha^i \alpha^k$ as given by (19.16). Remembering the non-relativistic theory, the first part $x_i \pi_k - x_k \pi_i$ of the angular momentum operator can be interpreted as the "orbital angular momentum" and the second part $\frac{\hbar}{2i} \alpha^i \alpha^k$ as the "spin angular momentum". It must, however, be kept in mind that in the relativistic theory this dichotomy of the angular momentum does not correspond to any directly observable physical fact. In a centrally symmetric electric field the torque d_{ik} evidently vanishes and the angular momentum remains constant in time.

22. Approximations of the Dirac Theory: The Non-Relativistic Quantum Mechanics of Spin as First Approximation

To make the transition to the non-relativistic theory of spin as a first approximation of the relativistic theory for slowly moving particles, it is useful to take representations for the matrices α^k and β given by eq. (18.19), where β is diagonal. Two of the components then turn out to be small compared to the other two if the velocity of the particle is small compared to the velocity of light. For demonstrating this, we introduce two two-component quantities (φ_1, φ_2) and (χ_1, χ_2) instead of the four-component quantity $(\psi_1, \psi_2, \psi_3, \psi_4)$ in which the factor $e^{-(i/\hbar)mc^2 t}$ will be separated, as is usual in non-relativistic quantum mechanics [cf. after eqs. (3.36) and (4.27'')]. Thus we have

$$\begin{aligned} \psi_1 &= \varphi_1 e^{-\frac{i}{\hbar} mc^2 t}, & \psi_2 &= \varphi_2 e^{-\frac{i}{\hbar} mc^2 t}, \\ \psi_3 &= \chi_1 e^{-\frac{i}{\hbar} mc^2 t}, & \psi_4 &= \chi_2 e^{-\frac{i}{\hbar} mc^2 t}. \end{aligned} \quad (22.1)$$

Then from (21.4):

$$\frac{\hbar}{i} \frac{\partial \varphi}{\partial t} - e \Phi_0 \varphi + \sum_{k=1}^3 c \sigma_k \left(\frac{\hbar}{i} \frac{\partial \chi}{\partial x_k} + \frac{e}{c} \Phi_k \chi \right) = 0, \quad (22.2a)$$

$$- 2mc^2 \chi + \frac{\hbar}{i} \frac{\partial \chi}{\partial t} - e \Phi_0 \chi + \sum_{k=1}^3 c \sigma_k \left(\frac{\hbar}{i} \frac{\partial \varphi}{\partial x_k} + \frac{e}{c} \Phi_k \varphi \right) = 0, \quad (22.2b)$$

where again the σ_k are the 2×2 matrices defined by eq. (18.17). Here it is essential that the term arising from the differentiation of the exponential factor with respect to the time cancels with the mass term multiplied by β in the case of the quantities φ . But the two add up for the quantities χ . This gives rise to the possibility of expansion in powers of the reciprocal of the velocity of light, $1/c$. If the quantities φ are considered as of the zeroth order, the quantities χ will be of the first order. If we introduce, as in (21.6), the operators

$$\pi_k = \frac{\hbar}{i} \frac{\partial}{\partial x_k} + \frac{e}{c} \Phi_k$$

we obtain up to quantities of the first order

$$\chi = \frac{1}{2mc} \sum_k \sigma_k \pi_k \varphi, \quad (22.3a)$$

and to a further approximation we will have

$$\chi = \frac{1}{2mc} \sum_k \sigma_k \pi_k \varphi + \frac{1}{4m^2c^3} \left(\frac{\hbar}{i} \frac{\partial}{\partial t} - e \Phi_0 \right) \sum_k \sigma_k \pi_k \varphi. \quad (22.3b)$$

If this is substituted in (22.2a), we get on retaining all quantities up to order $1/c^2$ inclusive:

$$\begin{aligned} \frac{\hbar}{i} \frac{\partial \varphi}{\partial t} - e \Phi_0 \varphi &+ \frac{1}{2m} \sum_k \sum_l \sigma_k \sigma_l \pi_k \pi_l \varphi \\ &+ \frac{1}{4m^2c^2} \sum_k \sum_l \sigma_k \pi_k \left(\frac{\hbar}{i} \frac{\partial}{\partial t} - e \Phi_0 \right) \sigma_l \pi_l \varphi = 0. \end{aligned}$$

By separating the terms with $k = l$ and $k \neq l$ and taking into consideration eqs. (21.7) and (18.18), it further follows that

$$\left. \begin{aligned} \left(1 + \frac{1}{4m^2c^2} \sum_k \pi_k^2 \right) \left(\frac{\hbar}{i} \frac{\partial \varphi}{\partial t} - e \Phi_0 \varphi \right) + \left\{ \frac{1}{2m} \sum_k \pi_k^2 + \frac{e\hbar}{2mc} \sum_k (\mathcal{H}_i \sigma_i) + \right. \right. \\ \left. \left. + \frac{e\hbar}{2mc} \frac{1}{2} \frac{1}{mc} \left(\sum_i [\vec{\sigma} \times \vec{\pi}]_i \sigma_i \right) - i \frac{e\hbar}{2mc} \frac{1}{2} \frac{1}{mc} \sum_i \mathcal{E}_i \pi_i \right\} \varphi = 0. \end{aligned} \right\} \quad (22.4)$$

The factor in (22.4) with which $\frac{\hbar}{i} \cdot \frac{\partial \varphi}{\partial t}$ is multiplied corresponds to the correction due to the variation of mass. The second term in the braces corresponds to the non-relativistic spin term in an external magnetic field with the correct value $-\frac{e\hbar}{2mc}$ of the magnetic spin moment; the third term inside the braces corresponds to the Thomas correction in an external electric field with the correct factor $\frac{1}{2}$ and finally the peculiar last term* which was first given by Darwin.¹

¹ C.G. Darwin, Proc. Roy. Soc. Lond., Ser. A 118, 654 (1928).

*Representing an imaginary electric moment.

This wave equation can also be obtained by inserting (22.3a) into the exact wave equation of the second order (21.8).

The above result contains the proof that for

$$\left| \left(\frac{\hbar}{i} \frac{\partial}{\partial t} - e \Phi_0 \right) \psi \right| \ll 2m c^2 |\psi| \quad (22.5)$$

the wave equation of non-relativistic quantum mechanics with spin follows from the Dirac wave equation as a first approximation. This approximation is enough, e.g., if one is dealing with a comparison of the eigenvalues of energy in the two theories. It is, however, essential that the consequences regarding the magnitude of the transition probabilities for the emission of light are in accord, within this approximation, in both the theories. This question will lead us back, according to the semiclassical treatment of the radiation process, to a comparison of the expressions for the current vector in the two theories.

To carry out this comparison, it is useful, to transform the current vector

$$s_\mu = i \psi^\dagger \gamma^\mu \psi$$

in a way given by Gordon.² According to the wave equations (III), (III†), we here replace first

$$\psi \quad \text{by} \quad - \frac{i}{mc} \sum_v \gamma^v \left(p_v + \frac{e}{c} \Phi_v \right) \psi,$$

and next

$$\psi^\dagger \quad \text{by} \quad + \frac{i}{mc} \sum_v \left[\left(p_v - \frac{e}{c} \Phi_v \right) \psi^\dagger \right] \gamma^v.$$

By separating the terms with $\mu = v$ and $\mu \neq v$, we obtain

$$s_\mu = s_\mu^{(0)} + s_\mu^{(1)}, \quad (22.6)$$

where

$$s_\mu^{(0)} = - \frac{1}{2m_0 c} \left\{ \left[\left(p_\mu - \frac{e}{c} \Phi_\mu \right) \psi^\dagger \right] \psi - \psi^\dagger \left(p_\mu + \frac{e}{c} \Phi_\mu \right) \psi \right\}, \quad (22.7)$$

$$s_\mu^{(1)} = \frac{\hbar}{2mc} \sum_v \frac{\partial M_{\mu v}}{\partial x_v}. \quad (22.8)$$

Here

$$M_{\mu v} = - M_{v \mu} = - i \psi^\dagger \gamma^\mu \gamma^v \psi \quad (22.9)$$

can be seen to be the tensor of polarisation and magnetisation. It is to be noted that

$$\sum_{\mu=1}^4 \frac{\partial s_\mu^{(1)}}{\partial x_\mu} = 0,$$

so that $s_k^{(0)}$ and $s_k^{(1)}$ separately satisfy the continuity equation. By changing over to ψ^* we obtain for the spatial components of the current density

$$i_k^{(0)} = c s_k^{(0)} = \frac{1}{2m_0} \left\{ \psi^* \left(p_k + \frac{e}{c} \Phi_k \right) \beta \psi - \left[\left(p_k - \frac{e}{c} \Phi_k \right) \psi^* \right] \beta \psi \right\}, \quad (22.7')$$

$$i_k^{(1)} = c s_k^{(1)} = \frac{\hbar}{2m} \sum_{v=1}^4 \frac{\partial M_{kv}}{\partial x_v} \quad (22.8')$$

and

$$\begin{aligned} M_{kl} &= - M_{lk} = \frac{1}{i} (\psi^* \beta \alpha^k \alpha^l \psi) \quad \text{for } k \neq l \quad \text{and} \quad k, l = 1, 2, 3 \\ M_{k4} &= (\psi^* \beta \alpha^k \psi). \end{aligned} \quad (22.9')$$

² W. Gordon, Z. Physik 50, 630 (1927).

If we again go over to the two-component notation according to (22.1) and specialise to the matrices α^k and β as given by (18.19), then we see that M_{kk} will be of the order $1/c^2$ and M_{kl} , apart from terms of order $1/c^2$, will be equal to

$$M_{12} = -M_{21} = (\varphi^* \sigma_3 \varphi), \dots \quad (22.10)$$

Ignoring terms of order $1/c^2$, $i_k^{[0]}$ in fact coincides with the expression for the current in the non-relativistic theory. The additional term

$$\vec{i} = \text{rot} (\varphi^* \vec{\sigma} \varphi) \quad (22.11)$$

does not, of course, give rise to a dipole radiation, according to eq. (15.20) and (15.23), since after carrying out the volume integration all its matrix elements vanish; but the term (22.11) has to be taken into account for the quadrupole and higher multipole radiations.

It is of interest to compare the expressions (21.16) and (21.19) for the angular momentum P_{ik} with the expression

$$\begin{aligned} \bar{M}_{ik} &= \frac{(-e)}{2c} \int (x_i i_k - x_k i_i) d^3 x \\ &= \int \psi^* \left\{ \frac{(-e)}{2mc} (x_i \pi_k - x_k \pi_i) \beta + \frac{(-e)\hbar}{2mc} \frac{1}{i} (\alpha^i \alpha^k \beta) \right\} \psi d^3 x \end{aligned} \quad \} \quad (22.12)$$

for the magnetic moment. On account of the appearance of the matrix β in the latter form, the two parts of \bar{M}_{ik} and P_{ik} are not in general proportional to each other. They are mutually proportional only for small velocities of the particle when quantities of the order v^2/c^2 can be neglected. In this case, the ratio of the magnetic and mechanical moments is equal to $-e/2mc$ for the first part and $-e/mc$ for the second part, as is required by experiments.³

For the semi-classical treatment of the scattering of radiation, Waller⁴ has compared the results from the Dirac wave equation with those arising from the wave equation of the non-relativistic theory. In the first wave equation, the perturbation Hamiltonian is simply $\sum_{k=1}^3 e(\alpha^k \Phi_k)$ if the vector potential of the external radiation field is substituted for Φ_k ; on the other hand, no terms proportional to Φ_k^2 are present in the perturbation, in contrast to the non-relativistic theory. Since in the latter theory, as mentioned in Sec. 15, precisely these terms in Φ_k^2 of the perturbation give the main contribution to the scattered radiation for an energy $h\nu$ of the incident radiation which is large compared to the ionisation potential of the system and small compared to mc^2 , one may, at first sight, doubt whether the results of the Dirac wave equation agree with those of the non-relativistic theory even approximately. It turns out, however, that those matrix elements of $\sum_k \alpha^k \Phi_k$ which correspond to transitions

from states of positive to states of negative energy are precisely the ones which give the intensity of the scattered radiation which, in the non-relativistic theory arise from the part of the Hamiltonian proportional to $\sum_k \Phi_k^2$. This is particularly of importance, since from this it has to be concluded that the matrix elements of the perturbation

³ Cf. for this also C.G. Darwin, Proc. Roy. Soc. Lond. **120**, 621 (1928); for the magnitude of the magnetic moment in hydrogen-like atoms, see G. Breit, Nature, Lond. **122**, 649 (1928).

⁴ I. Waller, Z. Physik **58**, 75 (1929).

which belong to the transitions mentioned above cannot simply be ignored. In particular these matrix elements prove to be essential for the agreement between the results on the intensity of the scattering of radiation by free electrons in the case $h\nu \ll mc^2$, as it follows on the one hand from the Dirac equation and from the classical theory (Thomson formula) on the other.

23. Approximations of the Dirac Theory: Limiting Transition to the Classical, Relativistic Particle Mechanics

A relativistic quantum theory must go over into known theories in two limiting cases, namely, the non-relativistic quantum mechanics on the one hand and the classical relativistic mechanics of particles on the other. Both these limiting cases can be roughly characterised as $\lim c \rightarrow \infty$ on the one hand and $\lim \hbar \rightarrow 0$ on the other. The first case was already dealt with in the previous section and the second will be considered now. The latter case is important, for example, for the discussion of the deflection experiments of electrons with velocities which are comparable to that of light, in external electric and magnetic fields; as is well known, such experiments serve to estimate the dependence of the mass of a particle on the velocity.

The classical relativistic mechanics of a particle of charge ($-e$) and rest mass m is built on the canonical equations of motion, arising from the Hamiltonian

$$H(p_k, x_k) = c \sqrt{m^2 c^2 + \sum_{k=1}^3 (p_k + \frac{e}{c} \Phi_k)^2} - e \Phi_0 \quad (23.1)$$

namely,

$$\dot{x}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial x_k}. \quad (23.2)$$

On introducing the quantities

$$\pi_k = p_k + \frac{e}{c} \Phi_k$$

these can also be written as

$$\dot{x}_k = \frac{c \pi_k}{\sqrt{m^2 c^2 + \sum_{k=1}^3 \pi_k^2}}, \quad \dot{\pi}_k = (-e) \left(\mathcal{E}_k + \frac{1}{2} [\vec{x} \times \vec{\mathcal{H}}]_k \right), \quad (23.3)$$

where $\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ are the field strengths derived from the potentials Φ_0 and Φ_k .

In order to investigate as to how far these statements follow as limiting cases from the wave equation, we must make a limiting transition from the wave equation to geometrical optics which is analogous to what was done in Sec. 12 for the case of non-relativistic quantum mechanics. Similar to the relation (12.1), we make here the ansatz

$$\psi_a = a_a e^{\frac{i}{\hbar} S} \quad (23.4)$$

and expand a_a in powers of \hbar/i :

$$a_a = a_{0a} + \frac{\hbar}{i} a_{1a} + \dots \quad (23.5)$$

It is essential here that the wave function (the eikonal) S does not depend on the in-

dex ρ since otherwise, on inserting it into the wave equation, the exponential factor cannot be removed and hence a meaningful expansion in powers of \hbar would be impossible. We obtain from the wave equation (21.4), on inserting (23.4) in it and with

$$\pi_0 = -\frac{1}{c} \frac{\partial S}{\partial t} + \frac{e}{c} \Phi_0, \quad \pi_k = \frac{\partial S}{\partial x_k} + \frac{e}{c} \Phi_k, \quad (23.6)$$

the equation

$$\left(-\pi_0 + \sum_{k=1}^3 \alpha^k \pi_k + \beta m c \right) a + \frac{\hbar}{i} \left(\frac{1}{c} \frac{\partial a}{\partial t} + \sum_{k=1}^3 \alpha^k \frac{\partial a}{\partial x_k} \right) = 0. \quad (23.7)$$

Here the indices have been again dropped so that a is to be understood as a one-column vector just like ψ (in contrast to S, π_0, π_k). By substituting the expansion (23.5) for a in eq. (23.7) we obtain, successively, the equations

$$\left(-\pi_0 + \sum_k \pi_k \alpha^k + m c \beta \right) a_0 = 0, \quad (23.8)$$

$$\left(-\pi_0 + \sum_k \pi_k \alpha^k + m c \beta \right) a_1 = - \left(\frac{1}{c} \frac{\partial a_0}{\partial t} + \sum_{k=1}^3 \alpha^k \frac{\partial a_0}{\partial x_k} \right), \quad (23.9)$$

In order that the system of homogeneous equations (23.8) possesses solutions, the π_0 and π_k must satisfy the condition

$$-\pi_0^2 + \sum_{k=1}^3 \pi_k^2 + m^2 c^2 = 0 \quad (23.10)$$

which, on account of (23.6) is identical with the Hamilton-Jacobi partial differential equation of particle mechanics. The discussion regarding eq. (23.9) then gives¹ the result that in the regions which we can reach, according to classical mechanics, i.e. where π_0 and π_k are real, the continuity equation

$$\frac{\partial \rho}{\partial t} + \sum_{k=1}^3 \frac{\partial}{\partial x_k} \left(\rho \frac{c \pi_k}{\pi_0} \right) = 0$$

holds for

$$\rho = (a_0^* a_0).$$

On account of

$$\frac{\partial H}{\partial p_k} = \frac{c \pi_k}{\pi_0}$$

and as a consequence of (23.10), this signifies that the particles themselves move along the classical trajectories defined by (23.2). This conclusion is quite analogous to that of the non-relativistic quantum mechanics. Here also the condition that a_1 is small compared to a_0 is satisfied by requiring that the gradient of the wave length of the matter waves should be numerically small.

A circumstance which is characteristic of the relativistic quantum mechanics is, however, that the resulting orbits are those of a particle without spin. The effects arising from spin on the space-time behaviour of the density and current of the particle become noticeable only in the amplitudes a_1 which also contain diffraction effects. In this second approximation, however, the concept of an orbit fails. This is due to the fact that the magnetic spin moment is proportional to the quantum of action and the

¹ For details cf. W. Pauli, Helv. Phys. Acta 5, 179 (1932).

effects arising from spin in the Dirac theory can be described without introducing an additional term.

This confirms the thesis of Bohr:² *The spin moment of the electron can never be determined, unambiguously separated from the orbital moment, by experiments to which the classical concept of the trajectory of a particle can be applied.* In fact, the discussion of any experiment for determining the spin moment of the free electron by deflection in suitable external force fields (e.g., by an arrangement similar to the Stern-Gerlach experiment for molecular beams) shows the following typical state of affairs: In order that the deviation is not swamped by diffraction effects, the beam must have sufficiently large dimensions. But then the effect of the Lorentz force arising from the variation of the field strength inside the beam makes it impossible to observe the deviation which will be due to the force acting on the spin alone.³

On the other hand, other experiments demonstrating the spin of the free electron are possible which have no relation to classical mechanics and to the concept of a trajectory. In this regard the possibility of a demonstration of the *polarisation of the electron waves* is of great interest. In analogy with the well-known experiment in classical wave optics, by a double scattering of an electron beam from an atom (or reflection from a mirror), the intensity of the tertiary ray will depend not only on the values of the scattering angle (angle of reflection) but also on the values of the angle Φ which the plane containing the primary and secondary beams forms with the plane containing the secondary and tertiary rays. And, in fact, in contrast to classical optics, where the intensity J of the tertiary ray depends only on $\cos^2 \Phi$, in the case of electrons this intensity is linear in $\cos \Phi$ for a fixed scattering angle and hence is of the form

$$J = J_0(1 + \delta \cos \Phi).$$

For the scattering of an electron from a bare nucleus this effect has been evaluated by Mott.⁴

A different type of possible polarisation effects can be obtained by using oriented atoms.⁵ We shall not go into this more closely, since till now there has been no reliable experiment which permits a unique comparison with the theory.

24. Transitions to States of Negative Energy. Limitations of the Dirac Theory

We have already seen in the free-particle case, that the wave equation possesses in addition to eigensolutions with positive energy, also eigensolutions which belong to

² N. Bohr, Atomic Theory and the Description of Nature, Cambridge Univ. Press (1961); Introductory survey, p. 9; further, his Faraday lecture, J. Chem. Soc. 1932, 349, in particular pp. 367 and 368.

³ For a deeper discussion cf. N.F. Mott, Proc. Roy. Soc. Lond., Ser. A 124, 425 (1929); C.G. Darwin, Proc. Roy. Soc. Lond. 130, 632 (1930); in addition, the Proc. of the Solvay Conference 1930, report by W. Pauli on the magnetic electron.

⁴ N.F. Mott, Proc. Roy. Soc. Lond. Ser. A 124, 425 (1929).

⁵ Cf. A. Landé, Naturwiss. 17, 634 (1929); E. Fues and H. Hellmann, Phys. Z. 31, 465 (1930); N.F. Mott, Proc. Roy. Soc. Lond., Ser. A 125, 222 (1929); also the Solvay-report, loc. cit., footnote 3 above.

negative energy. In the presence of fields of force, this fact may lead to peculiar paradoxes, as was first shown by Klein.¹

If we consider the classical theory with external forces, we see that the partial differential equation (23.10) of Hamilton-Jacobi permits both the solutions

$$\pi_0 = +\sqrt{m^2 c^2 + \sum_k \pi_k^2}$$

and

$$\pi_0 = -\sqrt{m^2 c^2 + \sum_k \pi_k^2}.$$

The latter case corresponds to a negative kinetic energy of the particle and to the Hamiltonian

$$H = -c \sqrt{m^2 c^2 + \sum_k \left(p_k + \frac{e}{c} \Phi_k \right)^2} - e \Phi_0. \quad (24.1)$$

The equation of motion is obtained from (23.3) by also changing the sign of the square root so that with $\dot{\pi}_k$ remaining unchanged in sign, we have

$$\dot{x}_k = -\frac{c \pi_k}{\sqrt{m^2 c^2 + \sum_k \pi_k^2}}. \quad (24.2)$$

Therefore, in this case, the acceleration is directed opposite to the force. The regions which correspond to a particular value of the total energy (kinetic plus potential) according to (23.1) (positive kinetic energy) and the regions which correspond to the same value of the total energy according to (24.1) (negative kinetic energy), are always spatially separated by a finite region in between.

Let us now go over to quantum mechanics and consider the special case of a one-dimensional electric field ($\Phi_k = 0$, Φ_0 depends only on x) and a wave function depending only on x (motion in the x -direction). Further, let Φ_0 decrease continuously with increasing x . For a given total energy E , we have then to distinguish between three regions:

- (1) $x < a$, $mc^2 < E + e\Phi_0(x)$,
- (2) $a < x < b$, $-mc^2 < E + e\Phi_0(x) < mc^2$,
- (3) $b < x$, $E + e\Phi_0 < -mc^2$.

The point $x = a$ corresponds to the turning point of the classical trajectory lying in the region (1) of a particle with positive kinetic energy and the point $x = b$ which lies further to the right corresponds to the turning point of the trajectory lying in the region (3) of a particle with negative kinetic energy and with the same value E of the total energy. Region (2) cannot be reached classically as the momentum

$$p(x) = \sqrt{\frac{[E + e\Phi_0(x)]^2}{c^2} - m^2 c^2} \quad (24.3)$$

is imaginary there.

In quantum mechanics this intermediate region is not, however, completely impenetrable. Under very general conditions, the transmission coefficient D is given by

$$D = e^{-2W}, \quad (24.4)$$

where

$$W = \frac{1}{\hbar} \int_a^b |p(x)| dx = \frac{1}{\hbar} \int_a^b \sqrt{m^2 c^2 - \left[\frac{E + e\Phi_0(x)}{c} \right]^2} dx. \quad (24.5)$$

¹ O. Klein, Z. Physik 53, 157 (1929).

W/\hbar denotes the action integral extended over the intermediate region.² Here it is assumed first that W is a very small number which, in practice, is always so and, secondly, that the potential Φ_0 is continuous. The original calculation of Klein relates to the singular case in which $\Phi_0(x)$ jumps discontinuously at a place so that the region (2) is compressed to a single point of the x -axis. In this case, in which the direct integration of the wave equation is easily done, the relation (24.4) fails.

The original interpretation of the Dirac equation in the framework of a one-body problem, therefore, leads to the result that particles with positive rest mass can with finite probability traverse the intermediate region and transform into particles with negative rest mass (preserving the value of the sum of the kinetic and potential energies). Evidently this result of the one-electron theory is in contradiction with experiments.

Dirac³ has successfully removed this difficulty by a new interpretation of his equation called "the hole theory". He imagines the vacuum to be described such that every electron state of negative energy is occupied by one electron. As a result of the exclusion principle this state is stable. He introduces further, the additional assumptions: (1) The infinite charge of these electrons does not produce a field; (2) Only the deviations in the occupation of the states from the normal occupation of the physical vacuum can produce an electrostatic field. In this case, the unoccupied states of negative energy behave like particles with charge $+e$ and positive mass with the same numerical value as that of the electron mass, both with respect to the field produced by them as also with regard to their behaviour in an external field. This antielectron or positron predicted by Dirac's theory has actually been found experimentally.*

Field quantisation enables us to give a more elegant formulation of the "hole theory" based on the exact symmetry of quantum electrodynamics (QED) with respect to the sign of the electrical charge (charge conjugation).

² Cf. W. Pauli, l.c. in footnote 1 of Sec. 23; for special potential-curves see F. Sauter, Z. Physik **69**, 742 (1931); **73**, 547 (1931). For a homogeneous field of strength F ,

$$W = \frac{\pi}{2} \frac{m^2 c^3}{\hbar e F}.$$

³ P.A.M. Dirac, Proc. Roy. Soc. Lond., Ser. A **126**, 360; **133**, 60 (1931). See The Principles of Quantum Mechanics, loc. cit. Refer further to J.R. Oppenheimer, Phys. Rev. **35**, 939 (1930).

*C.D. Anderson, Phys. Rev. **41**, 405 (1932).

CHAPTER X

Quantum Electrodynamics

25. Quantisation of the Free Radiation Field

(a) Classical Theory

As is well known, the eigen-vibrations of a cubical cavity, with edge l and volume $V = l^3$, whose walls are perfect reflectors, have the following properties: The components k_i of the wave vector whose magnitudes amount to 2π divided by the wave-length, have the eigenvalues

$$k_i = 2\pi \frac{s_i}{2l}, \quad s_1, s_2, s_3 = 1, 2, 3, \dots$$

where the integers s_i are to be restricted to positive numbers, since we are dealing with standing waves. The number dN of eigen-vibrations with wave vector components k_i lying between k_i and $k_i + dk_i$ is then given by

$$dN = V \cdot 2 \cdot 8 \cdot \frac{1}{(2\pi)^3} dk_1 dk_2 dk_3,$$

where the factor 2 takes care of the two polarisation directions of the wave, while the factor 8 arises because for standing waves we have to restrict ourselves to the positive octant of the k -space.

For the following considerations it is convenient to work with progressive waves. We obtain the same total number (dN) of eigen-vibrations in a given frequency interval, if we impose the following condition on the field: *The field must be periodic in each of the three space co-ordinates with period l .* Then the eigenvalues of k_i are

$$k_i = 2\pi \frac{s_i}{l} \quad \text{with} \quad s_i = 0, \pm 1, \pm 2, \dots \quad (25.1)$$

and can now be both positive and negative. The number of eigen-vibrations between k_i and $k_i + dk_i$ will be

$$dN = V \cdot 2 \frac{1}{(2\pi)^3} dk_1 dk_2 dk_3, \quad (25.2)$$

where the whole of k -space is taken into account and not merely the positive octant. The electric and magnetic field strengths $\vec{E}(k, x)$ and $\vec{H}(k, x)$ of the wave propagating along the direction of \vec{k} can conveniently be described, on introducing (one single) complex field strength vector $\vec{F}(k)$, in the following manner:

$$\vec{E}(k, x) = \vec{F}(k) e^{i\vec{k} \cdot \vec{x}} + \vec{F}^*(k) e^{-i\vec{k} \cdot \vec{x}} \quad (25.3)$$

$$\vec{H}(k, x) = \left| \frac{\vec{k}}{|k|}, \vec{E}(k, x) \right|. \quad (25.4)$$

Here we have the transversality condition

$$(\vec{k} \cdot \vec{E}(k, x)) = (\vec{k} \cdot \vec{H}(k, x)) = 0 \quad (25.5)$$

i.e.

$$(\vec{k} \cdot \vec{F}) = 0, \quad (\vec{k} \cdot \vec{F}^*) = 0. \quad (25.6)$$

There are no other relations between $\vec{F}(k)$, $\vec{F}(-k)$ and their complex conjugates. The decomposition of $\vec{E}(k, x)$ is done in such a way that the time dependence of F for all values of k is given by the factor $e^{-i\omega t}$ and that for F^* , by the factor $e^{+i\omega t}$:

$$\vec{F}(k, t) = \vec{F}(k, 0)e^{-i\omega t}, \quad \vec{F}^*(k, t) = \vec{F}^*(k, 0)e^{+i\omega t}. \quad (25.7)$$

Here ω denotes a positive number:

$$\omega = c |k|. \quad (25.8)$$

We, therefore, have for \vec{F} and \vec{F}^* , the differential equations:

$$\frac{d\vec{F}}{dt} = -ic |k| \vec{F}, \quad \frac{d\vec{F}^*}{dt} = +ic |k| \vec{F}^*. \quad (25.7')$$

Replacing \vec{k} by $-\vec{k}$ in (25.3) and (25.4), we obtain the field strengths of the waves propagating in the opposite direction.

Besides the decomposition of the field strength, $\vec{E}(k, x)$ into two parts with the time factors $e^{-i\omega t}$ and $e^{+i\omega t}$ for each k , it is often convenient to decompose the total field strength, spatially, as follows:

$$\vec{E}(k, x) + \vec{E}(-k, x) = \vec{E}(k)e^{\frac{i\vec{k} \cdot \vec{x}}{c}} + \vec{E}(-k)e^{-\frac{i\vec{k} \cdot \vec{x}}{c}}; \quad \vec{E}(-k) = \vec{E}^*(k) \quad (25.8a)$$

$$\vec{H}(k, x) + \vec{H}(-k, x) = \vec{H}(k)e^{\frac{i\vec{k} \cdot \vec{x}}{c}} + \vec{H}(-k)e^{-\frac{i\vec{k} \cdot \vec{x}}{c}}; \quad \vec{H}(-k) = \vec{H}^*(k). \quad (25.8b)$$

As is clear from a comparison with (25.3) and (25.4), we then have

$$\vec{E}(k) = \vec{F}(k) + \vec{F}^*(-k), \quad \vec{H}(k) = \left[\frac{\vec{k}}{|k|}, (\vec{F}(k) - \vec{F}^*(-k)) \right], \quad (25.9)$$

from which we have

$$\vec{F}(k) = \frac{1}{2} \left\{ \vec{E}(k) - \left[\frac{\vec{k}}{|k|}, \vec{H}(k) \right] \right\},$$

$$\vec{F}^*(-k) = \frac{1}{2} \left\{ \vec{E}(k) + \left[\frac{\vec{k}}{|k|}, \vec{H}(k) \right] \right\}. \quad (25.10)$$

The relations (25.6), (25.7) and (25.9) form complete expressions for the solution of the Maxwell equations. If we know, at a definite instant of time $\vec{E}(k, x)$ and $\vec{E}(-k, x)$, then $\vec{H}(k, x)$ is also determined, but on the other hand, the value of $\vec{H}(k)$ is not determined from (the values of) $\vec{E}(k)$ and $\vec{E}(-k)$ at this instant of time.

The energy $E(k)$ of the wave propagating in the direction of \vec{k} is given by

$$E(k) = \frac{1}{2} \int [\vec{E}^2(k, x) + \vec{H}^2(k, x)] dV = 2(\vec{F}(k) \vec{F}^*(k)) V, \quad (25.11)$$

on taking into account the relation $\vec{H}^2(k, x) = \vec{E}^2(k, x)$. The momentum of the wave is

$$\begin{aligned}\vec{P}(k) &= \int \frac{1}{2} [\vec{E}(k, x), \vec{H}(k, x)] dV = \frac{\vec{k}}{c|k|} \int \vec{E}^2(k, x) dV \\ &= \frac{\vec{k}}{c|k|} 2\vec{F}(k)\vec{F}^*(k)V.\end{aligned}\quad (25.12)$$

We can further decompose $\vec{F}(k)$ and $\vec{F}^*(k)$ and hence \vec{E} and \vec{H} , into two polarised eigen-vibrations. We set

$$\left. \begin{aligned}\vec{F}(k) &= \sum_{\lambda=1, 2} \vec{\epsilon}(\lambda, k) A(\lambda, k) = \sum_{\lambda=1, 2} \vec{\epsilon}(\lambda, k) B(\lambda, k) e^{i\omega t}, \\ \vec{F}^*(k) &= \sum_{\lambda=1, 2} \vec{\epsilon}^*(\lambda, k) A^*(\lambda, k) = \sum_{\lambda=1, 2} \vec{\epsilon}^*(\lambda, k) B^*(\lambda, k) e^{+i\omega t},\end{aligned}\right\} \quad (25.13)$$

where, we have on account of (25.7), for both values of λ ,

$$(\vec{\epsilon}(\lambda, k) \vec{k}) = 0, \quad \text{and} \quad (\vec{\epsilon}^*(\lambda, k) \vec{k}) = 0.$$

Also

$$\left. \begin{aligned}(\vec{\epsilon}(1, k) \vec{\epsilon}^*(2, k)) &= (\vec{\epsilon}^*(1, k) \vec{\epsilon}(2, k)) = 0, \\ (\vec{\epsilon}(1, k) \vec{\epsilon}^*(1, k)) &= (\vec{\epsilon}(2, k) \vec{\epsilon}^*(2, k)) = 1\end{aligned}\right\} \quad (25.14)$$

i.e. the $\vec{\epsilon}(1, k), \vec{\epsilon}(2, k)$ which are orthogonal to each other and to \vec{k} are in general complex unit vectors. This has the consequence that according to (25.10) the energy $E(k)$ is additively decomposed into

$$E(k) = 2 [A(1, k)A^*(1, k) + A(2, k)A^*(2, k)] V. \quad (25.14')$$

If, in particular, the three components of $\vec{\epsilon}$ are real numbers apart from a possible common phase factor, the radiation is linearly polarised.

Instead of considering the sum of field strengths, which correspond to the various eigenvalues \vec{k} , of \vec{k} , it is often convenient to take the limit $V \rightarrow \infty$. Then the periodicity condition drops out and we have to consider continuously varying \vec{k} and hence Fourier integrals. We have¹

$$\vec{E}(x) = \frac{1}{(2\pi)^{3/2}} \int \vec{E}(k) e^{i\vec{k} \cdot \vec{x}} dk^{[3]}, \quad \vec{E}(k) = \frac{1}{(2\pi)^{3/2}} \int \vec{E}(x) e^{-i\vec{k} \cdot \vec{x}} dx^{[3]} \quad (25.15)$$

and similarly for $\vec{H}(x)$ and $\vec{F}(x)$. Thus the relations between $\vec{E}(k), \vec{H}(k), \vec{F}(k)$ remain the same as given in (25.9) and (25.10). The total energy and total momentum of the wave-packet are then given by

$$E = \frac{1}{2} \int (\vec{E}^2 + \vec{H}^2) dx^{[3]} = 2 \int \vec{F}^*(k) \vec{F}(k) dk^{[3]}, \quad (25.16)$$

$$\vec{P} = \int \frac{1}{c} [\vec{E}, \vec{H}] dx^{[3]} = \frac{2}{c} \int \frac{\vec{k}}{|k|} \vec{F}^*(k) \vec{F}(k) dk^{[3]}. \quad (25.17)$$

Similarly, taking into account the transversality condition (25.6), we obtain for the angular momentum D with components $D_{ij} = -D_{ji}$ (on integration by parts) the

¹ We write $dk^{[3]}$ as an abbreviation for $dk_1 dk_2 dk_3$ and $dx^{[3]}$ as an abbreviation for $dx_1 dx_2 dx_3$.

expression:²

$$D_{ij} = \frac{1}{c} \int [\vec{x} \cdot (\vec{E}, \vec{H})]_{ij} dx^{[3]} = \frac{2i}{c} \int \frac{1}{|k|} \sum_{a=1}^3 \left(F_a^* \frac{\partial F_a}{\partial k_i} k_j - F_a^* \frac{\partial F_a}{\partial k_j} k_i \right) dk^{[3]} + \frac{2i}{c} \int \frac{1}{|k|} (F_j^* F_i - F_i^* F_j) dk^{[3]}. \quad (25.18)$$

We shall see later that this division of angular momentum into two parts is in some respects similar to the decomposition of the angular momentum of the electron into an orbital part and a spin part [cf. (21.19)].

(b) Quantisation

We now study the problem of the quantisation of the radiation field. For this purpose we start from the analogy between a polarised eigen-vibration and a harmonic oscillator. From experience with thermal equilibrium, it is known that the energy of such an eigen-vibration, with the wave vector \vec{k} , and the polarisation index λ ($=1, 2$) possesses the discrete eigenvalues:

$$E = N(\vec{k}, \lambda) \hbar \omega_r = N(\vec{k}, \lambda) \hbar c |\vec{k}_r|. \quad (25.19)$$

Here the fact that different eigen-vibrations are associated with independent quantum numbers has been taken account of. It is important to note that a zero-point energy of $\frac{1}{2}\hbar\omega_r$ per degree of freedom need not be introduced here, in contrast to the material oscillator. For, on the one hand, this would lead to an infinitely large energy per unit volume because of the infinite number of degrees of freedom of the system. On the other hand, this energy would, in principle, be unobservable, since it is neither emitted, nor absorbed nor scattered. So it cannot be enclosed by walls. Also, as is obvious from experience, it does not produce any gravitational field.

Earlier we applied the method of quantisation of the phase integral to the problem of the eigen-vibrations of cavity radiation, but now we apply the wave mechanical method. The latter method has the advantage that it can also be applied to the progressive waves. Evidently, we have only to express, according to eqs. (25.11) to (25.14), the fact that the energy

$$E_\lambda = 2V A^\dagger A_\lambda \quad (25.14')$$

of the polarised ($\lambda = 1, 2$) progressive wave possesses the eigenvalues $N \hbar \omega$. If we set

$$A_\lambda = \sqrt{\frac{\hbar \omega}{2V}} a_\lambda = \sqrt{\frac{\hbar c |k|}{2V}} a_\lambda; \quad A^\dagger = \sqrt{\frac{\hbar \omega}{2V}} a^\dagger = \sqrt{\frac{\hbar c |k|}{2V}} a^\dagger_\lambda, \quad (25.20)$$

then the operator

$$a^\dagger(\lambda, \vec{k}) a(\lambda, \vec{k}) = N(\lambda, \vec{k})$$

must have the eigenvalues 0, 1, 2, ... This is the case, if the quantities a^* and a (which are Hermitian conjugates of each other) satisfy the commutation relations

$$a(\lambda, \vec{k}_r) a^*(\lambda', \vec{k}_s) - a^*(\lambda', \vec{k}_s) a(\lambda, \vec{k}_r) = \begin{cases} 0 & \text{for } \lambda \neq \lambda' \text{ or } r \neq s \\ 1 & \text{for } \lambda = \lambda', r = s \end{cases} \quad (25.21)$$

while

$$a(\lambda, \vec{k}_r) a(\lambda', \vec{k}_s) - a(\lambda', \vec{k}_s) a(\lambda, \vec{k}_r) = 0$$

² Cf. C.G. Darwin, Proc. Roy. Soc. London (A), 136, 36 (1932).

and

$$a^*(\lambda, \vec{k}_r) a^*(\lambda', \vec{k}_s) - a^*(\lambda', \vec{k}_s) a^*(\lambda, \vec{k}_r) = 0. \quad (25.22)$$

[Cf. (14.22a) and (14.22a').] The zero-point energy of the oscillator is avoided by fixing the order of factors as has been done in eq. (25.14'), i.e. A^* standing before A . The symmetric expression

$$2V \frac{1}{2} (A^* A_\lambda + A_\lambda A^*) = \frac{\hbar\omega}{2} (a^* a_\lambda + a_\lambda a^*)$$

has the eigenvalues $(N + \frac{1}{2})\hbar\omega$ of the material harmonic oscillator. In fact the Hermitian quantities

$$p = \sqrt{\frac{\hbar\omega}{2}} (a + a^*), \quad q = i \sqrt{\frac{\hbar}{2\omega}} (a - a^*)$$

or

$$p = \sqrt{V} (A + A^*), \quad q = \frac{i}{\omega} \sqrt{V} (A - A^*),$$

which satisfy the commutation relations

$$pq - qp = -i\hbar,$$

are similar to the momentum and position of an oscillator. Even the expression for the energy, augmented by the zero-point energy, viz.,

$$E + \frac{\hbar\omega}{2} = V(A^* A + A A^*) = \frac{1}{2} (p^2 + \omega^2 q^2)$$

then takes the same form as for a harmonic oscillator with unit mass and the angular frequency ω . In the following it is convenient to calculate directly with the quantities a and a^* or A and A^* , instead of with p , q . The equations

$$a(\lambda, \vec{k}) = -ic|k| a(\lambda, \vec{k}); \quad a^*(\lambda, \vec{k}) = +ic|k| a^*(\lambda, \vec{k})$$

can be written, with the help of the Hamiltonian

$$H = \sum_{\lambda} \sum_{\vec{k}} 2VA^*(\lambda, \vec{k}_r) A(\lambda, \vec{k}_r) = \sum_{\lambda} \sum_{\vec{k}} \hbar c|\vec{k}_r| a^*(\lambda, \vec{k}_r) a(\lambda, \vec{k}_r) \quad (25.23)$$

in the usual form:

$$a(\lambda, \vec{k}_r) = \frac{i}{\hbar} [H, a(\lambda, \vec{k}_r)]; \quad a^*(\lambda, \vec{k}_r) = \frac{i}{\hbar} [H, a^*(\lambda, \vec{k}_r)]. \quad (25.24)$$

The operators a and a^* act on a wave function $\phi(N(k_r, \lambda))$ – indicated shortly by $\phi(N)$ for the case of a single eigen-vibration – in the following manner:

$$a^* \phi(N) = \sqrt{N} \phi(N-1), \quad a \phi(N) = \sqrt{N+1} \phi(N+1); \quad (25.25)$$

or in matrix form

$$a(N, N-1) = \sqrt{N}, \quad a^*(N-1, N) = \sqrt{N}. \quad (25.26)$$

The rest of the matrix elements vanish.

On introducing the auxiliary operators Δ^* and Δ , which satisfy the condition

$$\Delta^* \Delta = 1$$

and which are defined through

$$a^* = \sqrt{N} \Delta^*, \quad a = \Delta \sqrt{N}, \quad (25.27)$$

we obtain

$$\Delta \phi(N) = \phi(N+1), \quad \Delta^* \phi(N) = \phi(N-1) \quad (25.28)$$

[cf., for this, eqs. (14.24) to (14.27)].

By the above method of quantisation of the radiation field all corpuscular properties of light are already reproduced, e.g., the eigenvalue of the momentum of a progressive wave is, according to eqs. (25.12) and (25.19),

$$\vec{P}(\vec{k}) = N\hbar\vec{k}; \quad (25.19')$$

its magnitude is, therefore, $h\nu/c$. If we attribute a momentum $\hbar\vec{k}$ and an energy $h\nu$ to a photon, then the quantum number $N(\lambda, \vec{k})$ can be interpreted as *the number of photons with given momentum and given polarisation*. Even the fluctuations of energy and momentum of the radiation in a sub-volume can be described correctly through the formalism of the quantised waves, as was already mentioned in Sec. 15. Of course, these photons cannot be associated with the classical concept of a trajectory, *but the quantised waves are, in content, completely equivalent to the wave-mechanically described particles in their configuration or momentum space* (Sec. 15 and the following section).

Before proceeding further, we must speak about certain formal properties of the quantised waves. First, we shall indicate briefly how the functions of the photon numbers are transformed when we change from one kind of polarisation to another. This corresponds to a transformation of the amplitudes A in (25.13) into

$$A'(\lambda) = \sum_{\lambda' = 1, 2} c(\lambda, \lambda') A(\lambda'),$$

where c is unitary:

$$\sum c(\lambda, \lambda') c^*(\lambda'', \lambda') = \delta_{\lambda''} \text{. Hence } c(2, 1) = -c^*(1, 2); \quad c(2, 2) = c^*(1, 1).$$

To this corresponds a transition from the unit vectors $\vec{\epsilon}(1)$ and $\vec{\epsilon}(2)$ to the new unit vectors $\vec{\epsilon}'(1)$ and $\vec{\epsilon}'(2)$, according to

$$\vec{\epsilon}'(\lambda) = \sum_{\lambda'} c^*(\lambda', \lambda) \vec{\epsilon}(\lambda').$$

These vectors are again mutually orthogonal [eq. (25.14)]. If N'_1 and N'_2 denote respectively the number of photons corresponding to the vectors $\vec{\epsilon}'(1)$, $\vec{\epsilon}'(2)$ and $\phi'(N'_1, N'_2)$ is the associated eigenfunction, then according to the general transformation theory [see eq. (7.1)], the transition from the old photon numbers N_1 , N_2 and the associated eigenfunction, $\phi(N_1, N_2)$ to the new ones, is determined by the relation:

$$\phi'(N'_1, N'_2) = \sum_{N_1, N_2} \phi(N_1, N_2) S(N_1, N_2; N'_1, N'_2),$$

Here, according to eqs. (7.7) and (7.15), we must have, for all N_1 , N_2 , N'_1 , N'_2 and $\lambda = 1, 2$:

$$\begin{aligned} & \sum_{\bar{N}_1, \bar{N}_2} S(N_1, N_2; \bar{N}_1, \bar{N}_2) A'(\lambda) (\bar{N}'_1, \bar{N}'_2; N'_1, N'_2), \\ &= \sum_{\lambda' = 1, 2} c(\lambda, \lambda') \sum_{\bar{N}_1, \bar{N}_2} A(\lambda') (N_1, N_2; \bar{N}_1, \bar{N}_2) S(\bar{N}_1, \bar{N}_2; N'_1, N'_2). \end{aligned}$$

Substitution of the values (25.25) for the matrix elements of A [the constant factor which distinguishes $A(\lambda)$ from $a(\lambda)$ drops out] gives the recurrence formulae:

$$\begin{aligned} S(N_1, N_2; N'_1, N'_2) \cdot \sqrt{N'_1} &= c(1, 1) \sqrt{N_1} S(N_1 - 1, N_2; N'_1 - 1, N'_2) \\ &\quad + c(1, 2) \sqrt{N_1} S(N_1, N_2 - 1; N'_1 - 1, N'_2), \\ S(N_1, N_2; N'_1, N'_2) \cdot \sqrt{N'_2} &= c(2, 1) \sqrt{N_2} S(N_1 - 1, N_2; N'_1, N'_2 - 1) \\ &\quad + c(2, 2) \sqrt{N_2} S(N_1, N_2 - 1; N'_1, N'_2 - 1). \end{aligned} \quad (25.29)$$

We see easily that the $S(N_1, N_2; N'_1, N'_2)$ are different from zero only if $N_1 + N_2 = N'_1 + N'_2 = N$. Since S must be unitary, we further have $S(0, 0; 0, 0) = 1$. For a single photon we obtain from (25.29)

$$\begin{aligned} S(1, 0; 1, 0) &= c(1, 1), & S(0, 1; 1, 0) &= c(1, 2), \\ S(1, 0; 0, 1) &= c(2, 1), & S(0, 1; 0, 1) &= c(2, 2) \end{aligned}$$

and hence

$$\left. \begin{aligned} \phi'(1, 0) &= \phi(1, 0) c(1, 1) + \phi(0, 1) c(1, 2), \\ \phi'(0, 1) &= \phi(1, 0) c(2, 1) + \phi(0, 1) c(2, 2), \end{aligned} \right\} \quad (25.30)$$

i.e. ϕ transforms like $A(\lambda)$. In the general case of N photons, the transition from the $(N+1)$ -quantities $\phi(N_1, N-N_1)$, ($N_1 = 0, 1, \dots, N$) to the $N+1$ quantities $\phi'(N'_1, N-N'_1)$, ($N'_1 = 0, 1, \dots, N$) is similar to the transition from the $N+1$ expressions, $\sqrt{\frac{N}{N_1}} A_{(1)}^{N_1} A_{(2)}^{N-N_1}$ to the primed expressions $\sqrt{\frac{N}{N'_1}} A'_{(1)}^{N'_1} A'_{(2)}^{N-N'_1}$. For, both transformations determine an irreducible unitary representation of degree $N+1$ of the group of linear unitary transformations of two complex quantities (cf. Sec. 13).

We now proceed to set up the C.R. of the components of the vectors $\vec{E}(k)$, $\vec{H}(k)$ and $\vec{F}(k)$, first, for a definite value of \vec{k} . From eqs. (25.20), (25.21), (25.22), (25.13), (25.14) and from the relation

$$\sum_{\lambda=1,2} \epsilon_i(\lambda) \epsilon_j^*(\lambda) = \delta_{ij} - \frac{k_i k_j}{|k|^2}, \quad (25.14a)$$

which follows from (25.12) and the transversality condition, we obtain for $\vec{F}(k)$ the following relations:

$$\begin{aligned} [F_i(k), F_j(k')] &= 0, & [F_i^*(k), F_j^*(k')] &= 0, \\ [F_i(k), F_j^*(k)] &= [F_j(k), F_i^*(k)] = \frac{\hbar c |k|}{2V} \left(\delta_{ij} - \frac{k_i k_j}{|k|^2} \right). \end{aligned} \quad (25.31)$$

We shall from now onwards change over to a representation through continuous spectra. Then with the Hamiltonian operator

$$H = \int \hbar c |\vec{k}| \sum_{\lambda=1,2} a^*(\lambda, \vec{k}) a(\lambda, \vec{k}) dk^{[3]}, \quad (25.23')$$

(25.24) remains the same, but in the place of (25.21), we have to set

$$a(\lambda, \vec{k}) a^*(\lambda', \vec{k}') - a^*(\lambda', \vec{k}') a(\lambda, \vec{k}) = \delta_{\lambda\lambda'} \delta(\vec{k} - \vec{k}'),$$

where the second δ -function, which was introduced in (6.25) through

$$\int_V \delta(\vec{k}) dk^{[3]} = \begin{cases} 0, & \text{if } k = 0 \text{ outside } V_0, \\ 1, & \text{if } k = 0 \text{ inside } V_0 \end{cases}$$

defines an improper function. The relations (25.21) remain the same as also the first row of the relations, (25.31), but in the place of the second row of (25.31), we have

$$[E_i(k), E_j^*(k')] = [F_i(k), F_j^*(k')] = \frac{\hbar c |k|}{2} \left(\delta_{ij} - \frac{k_i k_j}{|k|^2} \right) \delta(\vec{k} - \vec{k}'). \quad (25.31')$$

The quantum number denoted earlier by $N(\lambda, \vec{k})$ goes over into the improper number $N(\lambda, \vec{k}) \delta(\vec{k} - \vec{k}')$ in the limiting transition to a continuous spectrum; we can still ask for the eigenvalues of

$$\int_{K_0} a^*(\lambda, k) a(\lambda, k) dk^{[3]} = N(\lambda, K_0), \quad (\lambda=1, 2) \quad (25.32a)$$

and

$$\int_{K_0} 2 \vec{F}^*(k) \vec{F}(k) dk^{[3]} = N(1, K_0) + N(2, K_0). \quad (25.34)$$

These are integers representing the number of photons in the interval of k -space indicated by K_0 corresponding to definite or indefinite polarisations. Of course, we have for the sum of two intervals K_1 and K_2 the identity, $N(\lambda, K_1) + N(\lambda, K_2) = N(\lambda, K_1 + K_2)$. The C.R. of the Fourier components E_i and $H_{ik} = -H_{ki}$ of the electric and magnetic field strengths (writing the latter for convenience as a skew-symmetric tensor) are obtained from the corresponding relations (25.31) and (25.31'), for the F_i , according to (25.9):

$$[E_i(k), E_j(k')] = 0, \quad [H_i(k), H_k(k')] = 0, \quad (25.33_1)$$

$$[E_i(k), H_j(k')] = \hbar c \delta(\vec{k} + \vec{k}') (\delta_{ij} k_i - \delta_{jk} k_j). \quad (25.33_2)$$

We note that here $\vec{k} + \vec{k}'$ occurs as the argument of the δ -function and not $\vec{k} - \vec{k}'$ and hence the point corresponding to $\vec{k} = -\vec{k}'$ is singular. We note further that the factors containing the magnitude of $|\vec{k}|$ have dropped out, in contrast to the C.R. for \vec{F} and \vec{F}^* . Moreover, the left-hand sides of the transversality conditions (25.5) and (25.6) commute with all quantities, as they should.

This facilitates the transition to the commutation relation for the field strengths written as functions of space:

$$\vec{E}(x) = \frac{1}{(2\pi)^{3/2}} \int \vec{E}(k) e^{i\vec{k}\vec{x}} dk^{[3]}, \quad \vec{H}(x) = \frac{1}{(2\pi)^{3/2}} \int \vec{H}(k) e^{i\vec{k}\vec{x}} dk^{[3]},$$

$$\vec{F}(x) = \frac{1}{(2\pi)^{3/2}} \int \vec{F}(k) e^{i\vec{k}\vec{x}} dk^{[3]}, \quad \vec{F}^*(x) = \frac{1}{(2\pi)^{3/2}} \int \vec{F}^*(k) e^{-i\vec{k}\vec{x}} dk^{[3]}.$$

Since we can formally set

$$\frac{1}{(2\pi)^3} \int e^{i\vec{k}\vec{x}} dk^{[3]} = \delta(\vec{x})$$

(which acquires proper meaning only after integration over a finite region of x -space, in the integrands on the left-hand sides) and hence

$$\frac{1}{(2\pi)^3} \int k_i e^{i\vec{k}\vec{x}} dk^{[3]} = \frac{1}{i} \frac{\partial}{\partial x_i} \delta(\vec{x}),$$

we obtain immediately for the field strengths, the C.R.³

³ These relations can be found in a somewhat different form (four-dimension) in P. Jordan and W. Pauli, Z. Phys. 47, 151 (1927) and in the form given here in W. Heisenberg and W. Pauli, ibid., 56, 1 (1927).

$$[E_i(x), E_j(x')] = 0, \quad [H_{ij}(x), H_{kl}(x')] = 0, \quad (25.34_1)$$

$$[E_i(x), H_j(x')] = \frac{\hbar c}{l} \left(\delta_{ij} \frac{\partial}{\partial x_l} - \delta_{il} \frac{\partial}{\partial x_j} \right) \delta(\vec{x} - \vec{x}').$$

For a corresponding formulation of the C.R. for the components of \vec{F} and \vec{F}^* we must define an operator $\sqrt{-\Delta}$ and its inverse $1/\sqrt{-\Delta}$. These are linear operators, which acting on the function $e^{i\vec{k}\vec{x}}$ give:

$$\sqrt{-\Delta} e^{i\vec{k}\vec{x}} = |\vec{k}| e^{i\vec{k}\vec{x}}, \quad (25.35_1)$$

$$\frac{1}{\sqrt{-\Delta}} e^{i\vec{k}\vec{x}} = \frac{1}{|\vec{k}|} e^{i\vec{k}\vec{x}}. \quad (25.35_2)$$

We see easily that the above two operators are Hermitian. Application of $\sqrt{-\Delta}$ twice gives the negative Laplacian operator $-\Delta$; hence the notation Δ . Thus the action of the operators $\sqrt{-\Delta}$ and $1/\sqrt{-\Delta}$ on any arbitrary function $f(x)$ is implicitly defined. We introduce the functions:

$$D_+(\vec{x}) = \sqrt{-\Delta} \delta(\vec{x}) = \frac{1}{(2\pi)^3} \int |\vec{k}| e^{i\vec{k}\vec{x}} dk^{(3)}, \quad (25.36_1)$$

$$D_{-}(\vec{x}) = \frac{1}{\sqrt{-\Delta}} \delta(\vec{x}) = \frac{1}{(2\pi)^3} \int \frac{1}{|\vec{k}|} e^{i\vec{k}\vec{x}} dk^{(3)} = \frac{1}{(2\pi)^2 r^2}. \quad (25.36_2)$$

The first is an improper function, for which only the integral over a finite region of x -space exists. Then⁴

$$\sqrt{-\Delta} f(\vec{x}) = \int f(\vec{x}') D_+(\vec{x} - \vec{x}') dx'^{(3)} \quad (25.37_1)$$

$$\frac{1}{\sqrt{-\Delta}} f(\vec{x}) = \int f(\vec{x}') D_{-}(\vec{x} - \vec{x}') dx'^{(3)} \quad (25.37_2)$$

and we have according to (25.31), (25.31')⁵

$$[F_i(x), F_j(x')] = 0, \quad [F_i^*(x), F_j^*(x')] = 0, \quad (25.38_1)$$

$$\begin{aligned} [F_i(x), F_j^*(x')] &= [F_i(x), F_i^*(x')] = \frac{1}{2} \hbar c \left(\sqrt{-\Delta} \delta_{ij} + \frac{1}{\sqrt{-\Delta}} \frac{\partial^2}{\partial x_i \partial x_j} \right) \delta(\vec{x} - \vec{x}') \\ &= \frac{1}{2} \hbar c \left\{ \delta_{ij} D_+ + \frac{\partial^2}{\partial x_i \partial x_j} D_{-}(\vec{x} - \vec{x}') \right\}. \end{aligned} \quad (25.38_2)$$

According to (25.16), the Hamiltonian operator will be

$$H = 2 \int \vec{F}^*(x) \vec{F}(x) dx^{(3)} \quad (25.39)$$

⁴ Cf. L. Landau and R. Peierls, Z. Phys. 62, 188 (1930).

⁵ A discussion on the introduction of the quantities F and F^* for avoiding the zero-point energy can be found in L. Rosenfeld and J. Solomon, Journ. de Phys. Series 7, 2, 139 (1931) and also in J. Solomon, Doctoral Thesis, Paris (1931). The C.R. between F and F^* given in these references are, incorrect, since they are not compatible with the conditions, $\operatorname{div} \vec{F} = 0$ and $\operatorname{div} \vec{F}^* = 0$.

and according to (25.17), the momentum is

$$\mathbf{P}_i = \frac{2}{c} \int \vec{F}^*(x) \frac{1}{i} \frac{\partial}{\partial x_i} \frac{1}{\sqrt{-\Delta}} \vec{F}(x) dx. \quad (25.40)$$

The transversality conditions are, simply,

$$\operatorname{div} \vec{F} = \operatorname{div} \vec{F}^* = 0. \quad (25.41)$$

These divergences commute with the Hamiltonian operator. Further,

$$\dot{\vec{F}} = \frac{i}{\hbar} [\mathbf{H}, \vec{F}] = -ic\sqrt{-\Delta} \vec{F}, \quad \dot{\vec{F}}^* = \frac{i}{\hbar} [\mathbf{H}, \vec{F}^*] = ic\sqrt{-\Delta} \vec{F}^*. \quad (25.42)$$

The occurrence of the operators $\sqrt{-\Delta}$ and $1/\sqrt{-\Delta}$ in the C.R. is hardly satisfactory, since these operators do not have infinitesimal (local) character, i.e. their value at a point depends on the whole spatial behaviour of the function and not merely on the behaviour of the function in the neighbourhood of the point considered. This again entails a quite non-perspicuous behaviour under Lorentz transformation of the quantities $\vec{F}(x)$ and $\vec{F}^*(x)$, which are connected with \vec{E} and \vec{H} , according to (25.9) and (25.10), through the relations,

$$\vec{E}(x) = \vec{F}(x) + \vec{F}^*(x), \quad \vec{H}(x) = \frac{1}{\sqrt{-\Delta}} \frac{1}{i} \operatorname{rot} (\vec{F}(x) - \vec{F}^*(x)), \quad (25.9')$$

$$\left. \begin{aligned} \vec{F}(x) &= \frac{1}{2} \left(\vec{E}(x) + \frac{i}{\sqrt{-\Delta}} \operatorname{rot} \vec{H} \right); \\ \vec{F}^*(x) &= \frac{1}{2} \left(\vec{E}(x) - \frac{i}{\sqrt{-\Delta}} \operatorname{rot} \vec{H} \right). \end{aligned} \right\} \quad (25.10')$$

The introduction of the quantities \vec{F} and \vec{F}^* which have been constructed in a rather artificial manner serves only to avoid the zero-point energy. The Hamiltonian operator will, on account of the relation

$$\frac{1}{2} \int (\vec{E}^2 + \vec{H}^2) dx^{[3]} = \int (\vec{F}^* \vec{F} + \vec{F} \vec{F}^*) dx^{[3]}$$

and (25.39) become

$$\mathbf{H} = \frac{1}{2} \int \{ (\vec{E}^2 + \vec{H}^2) + i[\vec{E}, -\frac{1}{\sqrt{-\Delta}} \operatorname{curl} \vec{H}] \} dx^{[3]}. \quad (25.43)$$

The expression in the bracket of the second term serves the purpose of subtracting the infinite zero-point energy which the first term contains; it includes likewise the operator $1/\sqrt{-\Delta}$.

The reason for the formal complications which occur here consists in the fact that the expectation values of functions of the field strengths at a definite space-point (e.g., quadratic functions) do not in general go over into the values of the classical quantities even in the limiting case of large quantum numbers and in fact will be infinitely large in many cases. This is due to the fact that here we are dealing with a system with infinitely many degrees of freedom (it is unimportant whether one assumes a countable infinity of degrees of freedom or a continuum of degrees of freedom). For example, the infinite product of eigenfunctions for the Fourier components of the electrical field strength does not converge, even if only a finite number of eigen-vibrations is excited. The use of the wave-mechanical formalism appears, therefore, to be justified on grounds of the correspondence principle if we

confine ourselves to a finite number of degrees of freedom. For instance, we can disregard sufficiently large eigenvalues of k in k -space; or in ordinary space, the construction of the mean of the field strengths over small but finite volumes may be carried out first, before the limiting transition to high quantum numbers (we shall see that this always happens for actual measurements of the field strengths). Alternatively, by using the number $N(k)$ of photons as variable (which is natural) we can confine ourselves to the case where only a finite number of photons is present. Applications of the theory to cases where a consideration of a finite number of degrees of freedom is not sufficient, lead to contradiction with experience (Sec. 8).

(c) *Limits of Accuracy on the Measurement of Field Strengths*⁶

It remains for us to investigate to what extent field strengths may be measured at all. The electrical field strength is defined through the change in the momentum of a test body with charge e , during a time δt , according to

$$e\vec{E} \cdot \delta t = \vec{P}' - \vec{P}.$$

If \vec{P} is known accurately before the measurement of the field strength and after time δt the momentum is measured with an accuracy $\Delta\vec{P}$ during the time Δt , then

$$e|\Delta\vec{E}|\delta t > \Delta\vec{P}. \quad (25.44)$$

Now for the momentum measurement of an arbitrary body, we have [eqs. (2.10), (2.11)]

$$\Delta P \Delta t > \frac{\hbar}{v - v'} > \frac{\hbar}{c} \quad (25.45)$$

It appears at first sight as if by using the test body with charge e the measurement of field strength can be made arbitrarily accurate. This statement is disputed by Landau and Peierls on the basis of the following argument. During the measurement of momentum, when the charged particle is accelerated, an amount of energy

$$\Delta E > \frac{e^2}{c^3} \frac{(v' - v)^2}{\Delta t}$$

is radiated.⁷ This gives an additional uncertainty in momentum,

$$\Delta P > \frac{\Delta E}{v' - v},$$

and hence

$$\Delta P \Delta t > \frac{e^2}{c^3} (v' - v). \quad (25.46)$$

But at this point there is an important lacuna in the argument of Landau and Peierls, since the radiated momentum and energy are amenable to an exact measurement. The implied change in the energy and momentum of the charged particle cannot, therefore, be considered directly as an *indeterminate* change. Hence there is an essential uncertainty in the further conclusions and the question regarding the accuracy of the measurement of field strengths must be treated as *not yet clarified*.

⁶ Cf. L. Landau and R. Peierls, Z. Phys. 69, 56 (1931) in particular, Secs. 3 and 4; W. Heisenberg, The Physical Principles of Quantum Theory, loc. cit., Chapter 3, Sec. 2.

⁷ We have $\int_{v'}^{v''} v^2 \Delta t > \frac{(v' - v)^2}{\Delta t}$ where Δt as well as the initial and final velocities are given.

From eqs. (25.45) and (25.46) follows, by multiplication,

$$\Delta P \Delta t > \frac{\hbar}{c} \sqrt{\frac{e^2}{\hbar c}}.$$

Hence

$$|\Delta \vec{E}| > \frac{\sqrt{\hbar c}}{(c \Delta t)^2}. \quad (25.47)$$

This inequality holds also for the magnetic field strengths:

$$|\Delta \vec{H}| > \frac{\sqrt{\hbar c}}{(c \Delta t)^2}. \quad (25.47')$$

This most favourable case occurs, if

$$\frac{e^2}{c^3} (v' - v)^2 \sim \hbar.$$

Since the mean frequency of the emerging light is $1/\Delta t$, this signifies that the mean number of emitted photons is at least of order 1. *The measurement of the field strength is connected with a finite but indefinite change in the number of photons.* The zero-point energy of those waves whose frequency v is smaller than $1/\Delta t$ corresponds precisely to the square of the field strength:

$$\vec{E}^2 \sim \frac{v^3 \hbar v}{c^3 2} \sim \frac{\hbar c}{(c \Delta t)^4},$$

which coincides with the square of the right-hand side of (25.47). Therefore this would not be measurable, if eq. (25.47) holds.

A further consideration shows that for simultaneous measurement of \vec{E} and \vec{H} in a spatial region Δl , we have

$$|\Delta \vec{E}| |\Delta \vec{H}| > \frac{\hbar c}{(c \delta t)^2} \cdot \frac{1}{(\Delta l)^2}. \quad (25.48)$$

This statement is sharper than the limiting value $\hbar c/(c \Delta t)^4$, which follows from multiplying (25.47) and (25.47'), only if $\Delta l < c \Delta t$. For wave fields, i.e. at distances from the bodies producing the field which are large compared to the wave-length, the inequality (25.48) yields nothing new. We have, further,

$$|\Delta \vec{E}| |\Delta \vec{H}| (\Delta l)^3 > \frac{\hbar c}{\Delta l} \quad \text{for} \quad \Delta t < \frac{\Delta l}{c} \quad (25.49)$$

which also follows directly from C.R. (25.34). Static fields can evidently be measured with arbitrary accuracy.⁸

⁸ The uncertainty relations (25.47) and (25.48) were considered independent of the question to which part of space a definite charge e can be compressed. For an electron, it follows from eqs. (25.44) and (25.45) without considering the emergent radiation, that

$$\Delta E > \frac{\hbar}{e c (\Delta t)^2} = \frac{\sqrt{\hbar c}}{e^2} \frac{\sqrt{\hbar c}}{(c \Delta t)^2}$$

which on account of the value $\hbar c/e^2 \sim 137$ is a higher limit than (25.47). The work of P. Jordan and V. Fock, Z. Phys., 66, 206 (1930) touches on the question of the accuracy of the field measurement with an electron. They find a somewhat different relation

$$|\Delta E| > \frac{\sqrt{\hbar c}}{e} \frac{\sqrt{\hbar c}}{c \Delta t / \Delta l}.$$

The circumstance that in the classical limiting case the field strengths \vec{E} and \vec{H} are measurable quantities as regards their space-time behaviour and hence also as regards their phases, has necessarily the consequence that the photons must correspond to symmetric states (i.e. obey the Bose-Einstein statistics). It is quite different for matter, where the wave functions ψ are not measurable quantities and where the case of the symmetric and antisymmetric states are, from the point of view of the correspondence principle, equally valid for many identical particles. Even a collection of material particles which have symmetric states, like, e.g., the He-nuclei, are not similar to an ensemble of photons, so long as no processes occur during which the number of particles changes. For, so long as this is not the case, there is no force similar to the Lorentz force for the ψ -function [considered as a q -number ψ -function in ordinary three-dimensional space, (cf. Sec. 14) and not as the c -number ψ -function in configuration space]. Its phase is included neither in the Hamiltonian operator nor in some other measurable physical quantity; the ψ -function cannot be measured.

(d) Transition to the Configuration Space of Photons⁹

Transition from the configuration space to the occupation number space, corresponding to the number of particles in unit volume of co-ordinate or momentum space (Sec. 14) is possible for photons just as in the case of identical material particles. If we assume that only one particle is present, then this connection is trivial. Let k_1, k_2, \dots be the values of k (assumed to be discrete). Then in the space of $N(\lambda, k)$, ($\lambda = 1, 2$), $\psi\{N(\lambda, k)\}$ is different from zero only if $N(\lambda, k_r)$ is equal to one for a certain position (k_r, λ_r) and zero otherwise; i.e. the eigenvalues of $N(\lambda, k_r)$ are

$$\delta_{\lambda_r} \delta(\vec{k}_r - \vec{k}_i).$$

If we substitute these eigenvalues for a definite λ_r and k_r in $\psi\{N(\lambda, k)\}$ as the argument and form according to eq. (25.13), the expression

$$\vec{f}(\vec{k}_i) = \vec{\epsilon}(\lambda_r, \vec{k}_i) \psi\{\delta_{\lambda_r} \delta(\vec{k}_r - \vec{k}_i)\},$$

then the vector \vec{f} perpendicular to \vec{k}_i can be considered to be the wave function of the photon in the momentum space. For N photons the eigenvalues of $N(\lambda, k_r)$ are

$$\sum_i \delta_{\lambda_r} \delta(\vec{k}_r - \vec{k}_i),$$

where we have summed over all the photon positions. Some of these points can be repeated. If p_1 denotes the number of single, p_2 the number of double, and p_N the number of N -fold points, such that

$$p_1 + 2p_2 + \dots + Np_N = N,$$

we have to form the combinatorial factor

$$C = \frac{N!}{(1!)^{p_1} (2!)^{p_2} \dots (N!)^{p_N}}$$

and set

$$\mathcal{F}_N(\vec{k}^{(1)}, \dots, \vec{k}^{(N)}) = \vec{\epsilon}^{(1)}(\lambda^{(1)}, \vec{k}^{(1)}) \dots \vec{\epsilon}^{(N)}(\lambda^{(N)}, \vec{k}^{(N)}) C^{-1/2} \psi\left\{\sum_i \delta_{\lambda_r} \delta(\vec{k}_r - \vec{k}_i)\right\}. \quad (25.50)$$

⁹ L. Landau and R. Peierls, Z. Phys., 62, 188 (1930); cf. also J.R. Oppenheimer, Phys. Rev., 38, 725 (1931).

\vec{f}_N is then a vector in $3N$ -dimensional space and hence has $3N$ components; it is perpendicular to all \vec{k}_i . The combinatorial factor is necessary in order that the functions \vec{f} are normalised, if $\phi(N(\lambda, k))$ is normalised. It is superfluous to go into the minute modifications that are necessary when k is interpreted as a continuous variable. The functions f are symmetric in the particle co-ordinates by definition, expressing the fact that the photons obey the Bose-Einstein statistics.

The application of the operators $F(k)$ on the wave functions \vec{f} is very easy because of eqs. (25.25) and (25.50). We have a series of functions

$$\vec{f}_0, \vec{f}_1(\vec{k}), \vec{f}_2(\vec{k}_1, \vec{k}_2), \dots, \vec{f}_N(\vec{k}_1, \vec{k}_2, \dots, \vec{k}_N), \dots$$

which relate to the various cases when zero, one, \dots , N photons are present. Then $F(k)$ transform the functions $f_N(\vec{k}_1, \dots, \vec{k}_N)$ into

$$F(k)f_N(i_1, \dots, i_N(\vec{k}_1, \dots, \vec{k}_N)) = f_{N+1}(i_1, \dots, i_N, i(\vec{k}_1, \dots, \vec{k}_N, \vec{k})). \quad (25.51)$$

The effect of F^* follows from the fact that it is conjugate to F and the results of applications of the field strength operators $E(k)$ and $H(k)$ follow from (25.9'). We shall see that the transformations, carried out under (b), on the wave function $\phi(N(\lambda, k))$ from one type of polarisation to another in configuration space is now a trivial transformation on f_N .

As an example, let us speak about the application to the angular momentum operators (25.18). They satisfy the same C.R. as the angular momentum operators of material particles [eq. (2.1)]. This must necessarily be the case, since it follows from consideration of the rotation group alone. Hence it has the same eigenvalues; each component D_{ij} has the eigenvalue m and the square, $D^2 = \sum_{i,j} D_{ij}^2$, has the eigenvalues $J(J+1)$. When one photon is present, the operator defined by (25.18) acts on $\vec{f}_i(\vec{k})$ in the following way:

$$D_{ij}f_i(k) = \frac{2i}{c|k|} \left(\left(\frac{\partial}{\partial k_i} k_j - k_i \frac{\partial}{\partial k_j} \right) f_i + (\delta_{ji}f_i - \delta_{ii}f_j) \right).$$

For a single photon we have the special rule that as a consequence of the transversality condition

$$\sum_i f_i k_i = 0$$

the eigenvalue $j=0$ is not allowed. In fact, for this case, we must have

$$\left(\frac{\partial}{\partial k_i} k_j - \frac{\partial}{\partial k_j} k_i \right) f_i + (\delta_{ji}f_i - \delta_{ii}f_j) = 0$$

for all i, j and i . We set $i=j$ and sum over j . Then

$$\sum_i \left[\frac{\partial}{\partial k_i} (k_i f_i) - k_i \frac{\partial f_i}{\partial k_i} \right] + 2f_i = 0.$$

This is, however, impossible on account of the transversality condition. For, we have immediately

$$2f_i = k_i \sum_i \frac{\partial f_i}{\partial k_i}$$

and then by scalar multiplication by \vec{k} , also

$$\sum_i \frac{\partial f_i}{\partial k_i} = 0.$$

Hence $f_i = 0$, i.e. all the components of \vec{f} must vanish. The selection rule established in Sec. 15 which forbids the transition $j = 0 \rightarrow j = 0$ for emission of a photon follows directly from the proof and the conservation law for angular momentum.

The functions $\vec{f}_N(k^{(1)} \dots k^{(N)})$ in momentum space determine corresponding functions $\vec{f}_N(x^{(1)} \dots x^{(N)})$ in the co-ordinate space, according to

$$\vec{f}_N(x^{(1)} \dots x^{(N)}) = \int \vec{f}_N(k^{(1)} \dots k^{(N)}) e^{i\vec{k}^{(1)} \cdot \vec{x}^{(1)} + \dots + i\vec{k}^{(N)} \cdot \vec{x}^{(N)}} dk^{(1)} \dots dk^{(N)} \quad (25.52)$$

These, however, have no direct relation to the particle density. For example, in the presence of one photon, $(\vec{f}^* \cdot \vec{x})(x)$ determines the energy density and not the spatial density of photons. The latter, we can try to define first through

$$(\vec{f}^* \frac{1}{\sqrt{-\Delta}} \vec{f})$$

or through

$$(\vec{g}^* \vec{g}),$$

where

$$\vec{g} = \frac{1}{\sqrt{-\Delta}} \vec{f}.$$

Here the $(\vec{g}^* \vec{g})$ is positive definite. But such a definition would be arbitrary and (as follows from the theory of the interaction of radiation and matter to be considered later) would not guarantee that the quanta would not act at the space-point where the density thus defined vanishes. *The vanishing of functions like \vec{f} or \vec{g} at a definite space-point has no direct physical significance.*

It turns out on account of the complicated behaviour of F under a Lorentz transformation that: *there is no current-density four-vector for a photon which satisfies the continuity equation and has a positive definite density.* Only one of the two requirements can be formally satisfied: either the vector character of the current-density under Lorentz transformation or the positive-definite character of the density. This stands in strict contrast to the description of material particles in Dirac's theory, in which, of course, both the requirements are satisfied. The non-existence of a density for the photon corresponds to the fact that the position of a photon cannot be associated with any operator in the usual sense (the position of the photon is not an "observable" in the sense of the definition of transformation theory, Sec. 7 and 9).

In fact the discussion of the possibility of a position measurement of a photon¹⁰ shows that the position cannot be determined more accurately than

$$|\Delta x| > \frac{hc}{E} \quad (25.53)$$

(E is the energy of the photon), and in an interval of time which is not shorter than

$$\Delta t > \frac{\hbar}{E} \quad (25.54)$$

This defines, however, precisely the *region of validity of geometrical optics*, since for a photon hc/E is equal to \hbar/P or equal to the wave-length. For a material particle the same inequalities hold: [eqs. (2.4)

¹⁰ L. Landau and R. Peierls, loc. cit.

and (2.5)], but hc/E can be essentially smaller than the wave-length of the matter wave. Only within the confines of the classical ray concept does the position of the photon have a physical significance.

The above statement will not apply to the measurement of the time averages (taken over time, which is large compared to a period of oscillation of the light wave) of \bar{E}^2 or \bar{H}^2 , which, e.g., for standing light waves can be measured even in space regions which are small compared to the wave-length of light.

26. Interaction Between Radiation and Matter

The theory of quantisation of the radiation field is complete only if it describes its interaction with matter. If n material particles are present, we shall denote the position and momentum co-ordinates of these particles (the momentum being divided by \hbar) by capital letters, like $\vec{X}_1 \dots \vec{X}_n$, and $\vec{K}_1 \dots \vec{K}_n$ in contrast to the corresponding lower case letters, $\vec{x}_1 \dots \vec{x}_N$ and $\vec{k}_1 \dots \vec{k}_N$ for the photons. Before we consider the question of choice of the Hamiltonian, we shall discuss the C.R. of the operators and the equations which describe the dependence of the operator on time. In order to have gauge-invariant quantities it is convenient to introduce for each of the n material particles, operators π_k ($k = 1, 2, 3$) which are analogous to those introduced in eq. (21.6), i.e.

$$\pi_k^{[s]} = p_k^{[s]} + \frac{e}{c} \Phi_k(X^{[s]}) = \frac{\hbar}{i} \frac{\partial}{\partial X_k} + \frac{e}{c} \Phi_k(X^{[s]}). \quad (26.1)$$

Here s is an index labelling the particles, which runs from 1 to n . The co-ordinates of the s^{th} particle are to be substituted in the potential. To start with, we shall ignore this definition of the operators $\pi_k^{[s]}$ and consider only their C.R., and their change with time. The C.R. are analogous to eq. (21.7):

$$\pi_i^{[s]} \pi_k^{[s']} - \pi_k^{[s']} \pi_i^{[s]} = \delta_{ss'} \frac{\hbar}{i} \frac{e}{c} H_{ik}(\vec{X}^{[s]}) \quad (26.2_1)$$

$$\pi_i^{[s]} X_k^{[s']} - X_k^{[s']} \pi_i^{[s]} = \delta_{ss'} \delta_{ik} \frac{\hbar}{i}. \quad (26.2_2)$$

The time dependence of the operators is analogous to eq. (21.15'):

$$\frac{d\pi_k^{[s]}}{dt} = (-e) \left\{ E_k(\vec{X}^{[s]}) + \sum_{l=1}^3 H_{kl}(\vec{X}^{[s]}) \alpha_l^s \right\}. \quad (26.3)$$

Then, according to Dirac's theory of the electron, the operators $\alpha_k^{[s]}, \beta^{[s]}$ introduced for each of the particles obey the relations:

$$\left. \begin{aligned} \alpha_i^{[s]} \alpha_k^{[s']} + \alpha_k^{[s']} \alpha_i^{[s]} &= 2\delta_{ss'} \delta_{ik}, & \alpha_i^{[s]} \beta^{[s']} + \beta^{[s']} \alpha_i^{[s]} &= 0 \\ [\alpha_i^{[s]}]^2 &= [\beta^{[s]}]^2 = 1 \end{aligned} \right\} \quad (26.4)$$

$$\vec{X}^{[s]} = c \alpha^{[s]}. \quad (26.5)$$

For the field strengths the C.R. (25.33) and (25.34) of vacuum (free-field)

electrodynamics are unaltered:

$$\left. \begin{aligned} [E_i(x), E_j(x')] &= 0, & [H_{ij}(x), H_{kl}(x')] &= 0, \\ [E_i(x), H_{jl}(x)] &= \frac{\hbar c}{i} \left(\delta_{ij} \frac{\partial}{\partial x_l} - \delta_{il} \frac{\partial}{\partial x_j} \right) \delta(\vec{x} - \vec{x}'), \end{aligned} \right\} \quad (25.34)$$

We further stipulate that the field strengths commute with the $X_k^{[s]}$ and $a_k^{[s]}$.

The most important requirement of the theory is that the field strength operators must now satisfy Maxwell's equations when charges are present:

$$\frac{1}{c} \frac{\partial \vec{H}}{\partial t} + \text{rot } \vec{E} = 0, \quad \text{div } \vec{H} = 0, \quad (26.6)$$

$$-\frac{1}{c} \frac{\partial \vec{E}(x)}{\partial t} + \text{rot } \vec{H}(x) = (-e) \sum_{s=1}^n \vec{a}^{[s]} \delta(\vec{x} - \vec{X}^{[s]}), \quad (26.7_1)$$

$$\text{div } \vec{E} = (-e) \sum_{s=1}^n \delta(\vec{x} - \vec{X}^{[s]}). \quad (26.7_2)$$

The appearance of the δ -function in the above equations corresponds to the assumption of a point charge in the classical theory. We shall later discuss the case when the δ -function is replaced by an arbitrary function $D(\vec{x} - \vec{X}^{[s]})$, which would correspond to a finite spread for the charges.

The last relation (26.7₂) is quite remarkable because it contains no time derivatives and hence must be satisfied at any instant of time; it represents a *subsidiary* condition. Its time derivative vanishes identically because of equations (26.7₁) and (26.5), which must necessarily be the case if the theory is to be consistent.

Now we have also to consider the C.R. of $\pi_k^{[s]}$ with the field strengths. These as well as the other C.R. must satisfy the following conditions: (1) they must be compatible with one another, (2) they evolve in time according to the time-dependent differential equations for the operators and (3) all operators must commute with the subsidiary condition (26.7₂).

This is the case, if we set¹

$$[\pi_i^{[s]}, H_{jl}(x)] = 0, \quad [\pi_i^{[s]}, E_j(x)] = (-e) \frac{\hbar}{i} \delta_{ij} \delta(\vec{x} - \vec{X}^{[s]}). \quad (26.8)$$

We then have

$$[\pi_i^{[s]}, \text{div } \vec{E}(x)] = (-e) \frac{\hbar}{i} \frac{\partial}{\partial x_j} \delta(\vec{x} - \vec{X}^{[s]})$$

in agreement with (26.7₂) and (26.2₂). Further the time derivatives of the C.R. (26.8) are satisfied on account of the remaining relations.

We then find an energy operator,

$$H = c \sum_{s=1}^n \left\{ \sum_{k=1}^3 a_k^{[s]} \pi_k^{[s]} + mc \beta^{[s]} \right\} + \frac{1}{2} \int (\vec{E}^2 + \vec{H}^2 + \Delta_0) dV \quad (26.9)$$

¹ In the literature these relations are derived from eq. (26.1) and the assumption $[\bullet_k(x), E_j(x')] = (\hbar/i) \delta(\vec{x} - \vec{x}')$ which, however, we shall not be using.

which commutes with the subsidiary condition (26.7₂) and a momentum operator

$$\mathbf{P}_i = \sum \pi_i^{[s]} + \frac{1}{c} \int \{ (\vec{E}, \vec{H})_i + \Delta_i \} dV. \quad (26.10)$$

The auxiliary quantities Δ_0, Δ_i in the integrands which we have not explicitly written down, commute with all the operators (i.e. they are *c*-numbers) and serve to remove the zero-point energy [cf. (25.43)]. For each of the operators

$$\pi_i^{[s]}, X_i^{[s]}, \alpha_i^{[s]}, \beta_i^{[s]}; \quad E_i(x), H_{jk}(x)$$

used and in general for each of the functions f of these operators (which do not contain the co-ordinates x_k and t explicitly), we have

$$\dot{f} = \frac{i}{\hbar} [\mathbf{H}, f], \quad (26.11)$$

$$\sum_s \frac{\partial f}{\partial X_k^{[s]}} + \frac{\partial f}{\partial x_k} = \frac{i}{\hbar} [\mathbf{P}_k, f]. \quad (26.12)$$

We could have also started from the requirement of the existence of the Hamiltonian and momentum operators and could have derived the temporal differential equations from them according to eq. (26.11). The absence of the scalar potential in eq. (26.9) is to be noted. The term $(-e) E_k$ on the right-hand side of eq. (26.3) follows directly from the C.R. of π_k with $(\vec{E})^2$.

We now study the relativistic invariance of the theory developed above. If we consider an orthogonal co-ordinate transformation

$$x'_\mu = \sum_{\nu=1}^4 a_{\mu\nu} x_\nu$$

($x_4 = ict$), we have operators in the new reference system; i.e. we go over from the set

$$\pi_i^{[s]}(t), X_i^{[s]}(t), \alpha_i^{[s]}(t), \beta_i^{[s]}(t); \quad E_i(x, t), H_{jk}(x, t)$$

to the set

$$\pi_i'^{[s]}(t'), X_i'^{[s]}(t'), \alpha_i'^{[s]}(t'), \beta_i'^{[s]}(t'); \quad E_i'(x', t'), H_{jk}'(x', t').$$

In quantities referring to a material particle we have to substitute for t' ,

$$ict'^{[s]} = a_{44}ict + \sum_{r=1}^3 a_{4r}X_r^{[s]}(t).$$

The field strengths must transform like a skew-symmetric tensor in four-dimensional space. The quantities γ_μ , namely $(-i\beta a_\mu, \beta)$, defined by (19.1) transform like a four-vector. The $X_{(n)}^i$ together with ict and $\{\pi_k, t(\sum a_i \pi_i + mc\beta)\}$ form again two four-vectors. Further $(\mathbf{P}_k, \frac{i}{c}\mathbf{H})$ must form a four-vector. Also the C.R. in the primed and unprimed systems must have the same form. To verify the relativistic invariance of the theory, it is convenient, instead of fixing the world point, to change it such that primed co-ordinates of the new world point have the same value as the unprimed co-ordinates of the old world point, i.e., we go over from the unprimed co-ordinates to

$$\pi_i'^{[s]}(t), X_i'^{[s]}(t), \alpha_i'^{[s]}(t), \beta_i'^{[s]}(t); \quad E_i'(x, t), H_{jk}'(x, t),$$

where t and (x, t) are the values in the primed co-ordinates. We shall call this a transformation of the second kind, while the one given earlier can be designated a transformation of the first kind. *The relativistic invariance of the theory is proved if a unitary operator S exists which effects for each of the quantities mentioned and hence also for every function f of them, a transformation of the second kind, according to*

$$f'(t) = S f(t) S^{-1}. \quad (26.13)$$

For proving this it is sufficient to show that this is true for infinitesimal co-ordinate transformations

$$x'_k = x_k + \sum_{\nu=1}^4 \varepsilon_{\mu\nu} x_\nu, \quad \varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu},$$

where only quantities of the first order in $\varepsilon_{\mu\nu}$ are retained. On setting

$$S = I + \frac{i}{\hbar} \Lambda, \quad (26.14)$$

we get

$$f'(x, t) = f(x, t) + \frac{i}{\hbar} [\Lambda, f(x, t)] \quad (26.15)$$

instead of eq. (26.13). The operator Λ depends linearly on $\varepsilon_{\mu\nu}$ according to

$$\Lambda = \sum_{\mu<\nu} \Lambda_{\mu\nu} \varepsilon_{\mu\nu}, \quad (26.16)$$

where $\Lambda_{\mu\nu} = -\Lambda_{\nu\mu}$ form, as we shall see, the components of an antisymmetric tensor, which are constant in time and hence are integrals of the field equations. The infinitesimal transformation of the first kind of the quantity f is obtained from the infinitesimal transformation of the second kind, in the following way. To a function f of the field strengths, at a fixed point of space, we have to add

$$\sum_{k=1}^3 \left\{ \sum_{\nu=1}^4 \frac{\partial f}{\partial x_k} \varepsilon_{k\nu} x_\nu + \frac{1}{ic} \int \varepsilon_{4k} x_k \right\}$$

and to $\dot{\pi}_j^{(s)}$ and $X_j^{(s)}$, we must add

$$\sum_{k=1}^3 X_k^{(s)} \varepsilon_{4k} X_k^{(s)}, \quad \sum_{k=1}^3 \dot{\pi}_j^{(s)} \varepsilon_{4k} X_k^{(s)}.$$

We shall now write the $\Lambda_{\mu\nu}$ for spatial rotations and Lorentz transformations separately. We have to set

$$\begin{aligned} \Delta_{jk} &= \sum_j (X_j^{(s)} \pi_k^{(s)} - X_k^{(s)} \pi_j^{(s)} + \frac{\hbar}{2} \frac{1}{i} a_j^{(s)} a_k^{(s)}) \\ &+ \frac{1}{c} \int (\vec{E} \cdot \vec{H})_{jk} dV, \end{aligned} \quad (26.17_1)$$

$$\begin{aligned} \frac{1}{i} \Lambda_{ij} &= c i P_i - \sum_s \left\{ X_j^{(s)} \left(\sum_{k=1}^3 a_k^{(s)} \pi_k^{(s)} + mc \beta^{(s)} \right) + \frac{\hbar}{2} a_j^{(s)} \right\} \\ &- \frac{1}{2c} \int x_j (\vec{E}^2 + \vec{H}^2 + \Delta_0) dV. \end{aligned} \quad (26.17_2)$$

P_j is again the component of momentum given by (26.10). We recognise in Λ_{jk} the angular momentum integral which is a sum of the part (21.19) corresponding to matter and (25.18) corresponding to radiation; further the Λ_{ij} are the additional space-time components, which together with Λ_{jk} form an antisymmetric tensor. As calculation shows the time derivatives of these quantities vanish so that they are

likewise integrals of the field equations (equations of motion). In the terms $\frac{\hbar}{2i} \alpha_j^{(l)} \partial_k^{(l)}, \frac{\hbar}{2} \alpha^{(l)}$ which can be combined into $\frac{1}{2} Y_\mu^{(l)} Y_\nu^{(l)}$, we recognise the quantities which are significant for the transformation of Dirac's wave functions (here the $\alpha_k^{(l)}, \beta^{(l)}$) according to eq. (25.22). The calculation of the commutators of \mathbf{A} with all the operators present shows that these transform correctly and that the momentum and energy form a four-vector; thus the proof of the relativistic invariance of the theory is complete.

There are already general systematic investigations on the quantisation of arbitrary classical field equations and the possibility of deriving the C.R. from a canonical scheme. We need not, therefore, go into them, nor even into their connection with the quantisation of matter waves spoken of in Sec. 15, but refer the reader to the literature.² Interesting features come to the fore when the Hamiltonian possesses invariance under a group, whose transformations contain arbitrary functions. In the case of quantum electrodynamics this group is the group of gauge transformations. In the representation given here we have used only gauge-invariant quantities, without using a systematic method of finding the C.R. If we wish to retain the potentials, the method of Fermi is perhaps the most transparent. According to this method, we set as the radiation part of the Hamiltonian operator

$$\frac{1}{2} \int \left\{ \sum_k \left[\sum_i \left(\frac{\partial \Phi_k}{\partial x_i} \right)^2 + \frac{1}{c^2} \left(\frac{\partial \Phi_k}{\partial t} \right)^2 \right] - \left[\sum_i \left(\frac{\partial \Phi_0}{\partial x_i} \right)^2 + \frac{1}{c^2} \left(\frac{\partial \Phi_0}{\partial t} \right)^2 \right] \right\} dV.$$

$\partial \Phi_k / \partial t$ and $\partial \Phi_0 / \partial t$ play the role of the momenta canonically conjugate to Φ_k and Φ_0 respectively; we have to add the term $-e \sum_i \Phi_0 (X^{(i)})$ to the matter part of the Hamiltonian. Then we must require the validity of the subsidiary condition

$$\sum_k \frac{\partial \Phi_k}{\partial x_k} + \frac{1}{c} \frac{\partial \Phi_0}{\partial t} = 0$$

and its time derivatives which can be brought to the form

$$\operatorname{div} \vec{E} = \rho = (-e) \sum_i \delta(x - X^{(i)}).$$

Both the conditions commute with the Hamiltonian because they are valid at any time once they are both satisfied at time $t = 0$. The subsidiary conditions indeed restrict the gauge transformations essentially, but in order to check the invariance of the theory under Lorentz transformations, Fermi's method is quite suitable. The results are identical with those given here or with any other consistent treatment of quantum electrodynamics.

In order to draw conclusions in a particular frame of reference from the fundamental equations of the theory, it is convenient to decompose the electric field strength \vec{E} into a longitudinal part and a transverse part:

$$\vec{E} = \vec{E}^{(l)} + \vec{E}^{(tr)}. \quad (26.18)$$

This equation means that by a Fourier spatial decomposition of \vec{E} , i.e. transition to k -space, $\vec{E}^{(l)}(k)$ will be parallel to \vec{k} while $\vec{E}^{(tr)}$ will be perpendicular to \vec{k} . In coordinate space this is equivalent to saying that $\vec{E}^{(l)}$ is irrotational and $\vec{E}^{(tr)}$ is solenoidal:

$$\operatorname{rot} \vec{E}^{(l)} = 0, \quad \operatorname{div} \vec{E}^{(tr)} = 0. \quad (26.19)$$

² Reviews: L. Rosenfeld, Mém. de l'Inst. Henri Poincaré, **2**, 24 (1932); E. Fermi, Rev. Mod. Phys. **1**, 87 (1932). Original Papers: G. Mie, Ann. d. Phys. (4) **85**, 711 (1928); W. Heisenberg and W. Pauli, I, Z. Phys. **56**, 1 (1929); II, ibid., **59**, 168 (1929). (Remarks on these: L. Rosenfeld, ibid., **58**, 540 (1929) and **63**, 574 (1930); E. Fermi, Lincei Rend. (6), **9**, 881 (1929), **12**, 431 (1930); L. Rosenfeld, Ann. d. Phys., **5**, 113 (1930).)

For $\vec{E}^{(1)}$ we have, from (26.7₂)

$$\vec{E}^{(1)} = -\text{grad}(-e) \sum_s \frac{1}{r_s} = -e \sum_s \frac{\vec{x} - \vec{x}_s}{r_s^3} \quad (26.20)$$

if

$$r_s = |\vec{x} - \vec{x}^{(s)}|$$

denotes the distance of the reference point from the position of the particle.³ We then have in a well-known manner

$$\int \vec{E}^{(1)} \vec{E}^{(1)} dV = 0,$$

$$\frac{1}{2} \int (\vec{E}^{(1)})^2 dV = \frac{1}{2} \sum_s \sum_{s'} \frac{e^2}{r_{ss'}}.$$

if

$$r_{ss'} = |\vec{x}^{(s)} - \vec{x}^{(s')}|$$

denotes the distance of the particle s from the particle s' . For $s = s'$ terms which are infinitely large are included; this sum is the infinite electrostatic self-energy of the particle (which evidently occurs even for the case of individual particles). In order to make calculations with the theory, we must replace

$$\frac{1}{2} \int (\vec{E}^{(1)})^2 dV = \frac{1}{2} \sum_s \sum_{s'} \frac{e^2}{r_{ss'}} = n \cdot \infty + \sum_{s < s'} \frac{e^2}{r_{ss'}} + \sum_{s > s'} \frac{e^2}{r_{ss'}}.$$

by

By this, however, the relativistic invariance of the theory is lost.

With this modification, the Hamiltonian operator (26.9) is transformed into

$$H' = c \sum_{s=1}^n \left\{ \sum_k \alpha_k^{(s)} \pi_k^{(s)} + mc \beta^{(s)} \right\} + \sum_{s < s'} \frac{e^2}{r_{ss'}} + \frac{1}{2} \int (\vec{E}^{(1)s} \vec{H}^2 + \vec{H}^2 + \Delta_0) dV. \quad (26.21)$$

Here, as in vacuum electrodynamics, $\vec{E}^{(1)s}$ can be expressed in terms of photon number operator; but now the number of photons will vary with time, while it was constant in vacuum electrodynamics.

We now introduce a Schrödinger wave function

$$\psi_{p_1 p_2 \dots p_n} (\vec{x}^{(1)}, \dots, \vec{x}^{(n)}; N(\lambda, \vec{k}))$$

which depends on the spin indices (each running from 1 to 4) and the co-ordinates of the material particles, on the one hand and the photon numbers in momentum space, on the other. The square of the absolute value of this function denotes the corresponding probability. The way in which the operators $\alpha_k^{(s)}, \beta^{(s)}, \vec{x}_k^{(s)}, N(\lambda, k)$ act on ψ is clear, but the effect of the field strengths \vec{E} and \vec{H} and the $\pi_k^{(s)}$ is not, on the contrary, uniquely determined. For, in definition of ψ a phase factor, $e^{i/(X^{(1)} \dots X^{(n)}; N(\lambda, k))}$ is arbitrary and according to the way this factor is chosen, the application of

³ We always assume that \vec{E} vanishes sufficiently fast at infinity. Then even fields which are both solenoidal as well as irrotational are excluded.

the operators \vec{E} , \vec{H} and $\pi_k^{[s]}$ on ψ will give different results. Any factor which is in agreement with C.R. is allowed.

We can first stipulate that \vec{E}^{tr} and \vec{H} shall act as in vacuum electrodynamics, i.e. according to eqs. (25.9), (25.13) and (25.20), we have

$$\begin{aligned}\vec{E}^{\text{tr}} &= \int \vec{E}(k) e^{i(\vec{k} \cdot \vec{x})} dk^{[3]} = \int (F(k) + F^*(-k)) e^{i(\vec{k} \cdot \vec{x})} dk^{[3]} \\ &= \int \sqrt{\frac{\hbar c |k|}{2}} \sum_{\lambda=1,2} [\vec{\epsilon}(\lambda, k) \mathbf{a}(\lambda, k) + \vec{\epsilon}^*(\lambda, -k) \mathbf{a}^*(\lambda, -k)] e^{i(\vec{k} \cdot \vec{x})} dk^{[3]}, \\ \vec{H} &= \int \vec{H}(k) e^{i(\vec{k} \cdot \vec{x})} dk^{[3]} = \int [\frac{\vec{k}}{|k|}, \vec{F}(k) - \vec{F}^*(-k)] e^{i(\vec{k} \cdot \vec{x})} dk^{[3]} \\ &= \int \sqrt{\frac{\hbar c |k|}{2}} \sum_{\lambda=1,2} \left\{ \left[\frac{\vec{k}}{|k|}, \vec{\epsilon}(\lambda, k) \right] \mathbf{a}(\lambda, k) \right. \\ &\quad \left. + \left[\frac{\vec{k}}{|k|}, \vec{\epsilon}^*(\lambda, -k) \right] \mathbf{a}^*(\lambda, -k) \right\} e^{i(\vec{k} \cdot \vec{x})} dk^{[3]}.\end{aligned}$$

The operators \mathbf{a} , \mathbf{a}^* act as given by eqs. (25.25) to (25.28).

We then obtain the following condition on the action of $\pi_j^{[s]}$ which is compatible with C.R. (26.2) and (26.8):

$$\begin{aligned}\pi_j^{[s]} \psi (\vec{X}, N(k)) &= \left\{ \frac{\hbar}{i} \frac{\partial}{\partial X_j^{[s]}} + \frac{e}{c} \int \frac{1}{r_s} \sum_{k=1}^3 \frac{\partial H_{jk}(x)}{\partial x_k} dV \right\} \psi \\ &= \left\{ \frac{\hbar}{i} \frac{\partial}{\partial X_j^{[s]}} + \frac{e}{c} \frac{-i}{\sqrt{-\Delta}} [\vec{F}(\vec{X}_s) - \vec{F}_j^*(\vec{X}_s)] \right\} \psi,\end{aligned}$$

or written in k -space,

$$\begin{aligned}\pi_j^{[s]} \psi (\vec{X}, N(k)) &= \left\{ \frac{\hbar}{i} \frac{\partial}{\partial X_j^{[s]}} + \frac{e}{c} (-i) \int \frac{1}{|k|} \right. \\ &\quad \times [F_j(k) - F_j^*(-k)] e^{i(\vec{k} \cdot \vec{X}^{[s]})} dk^{[3]} \Big\} \psi. \quad (26.22_1)\end{aligned}$$

Hence

$$\begin{aligned}\pi_j^{[s]} \psi (\vec{X}, N(k)) &= \left\{ \frac{\hbar}{i} \frac{\partial}{\partial X_j^{[s]}} + \frac{e}{c} (-i) \int \sqrt{\frac{\hbar c}{2|k|}} e^{i(\vec{k} \cdot \vec{X}^{[s]})} \right. \\ &\quad \left. [\epsilon_j(\lambda, k) \mathbf{a}(\lambda, k) + \epsilon_j^*(\lambda, -k) \mathbf{a}^*(\lambda, -k)] dk^{[3]} \right\} \psi. \quad (26.22_2)\end{aligned}$$

If we substitute this in the Hamiltonian operator, we obtain

$$\begin{aligned}\frac{\hbar}{i} \frac{\partial \psi}{\partial t} + \left\{ \int N(k) \hbar c |k| dk^{[3]} + c \sum_{s=1}^n \left(\sum_{k=1}^3 \mathbf{a}_k^{[s]} \frac{\hbar}{i} \frac{\partial}{\partial X_k^{[s]}} \right. \right. \\ \left. \left. + mc\beta^{[s]} + \sum_{s \leq s_0} \frac{e^2}{r_s} + H_1 \right) \right\} \psi = 0, \quad (26.23)\end{aligned}$$

where

$$H_1 \psi = e(-i) \left\{ \int \sum_{k=1}^3 \sqrt{\frac{\hbar c}{2|k|}} \cdot \sum_{s=1}^n e^{i(\vec{k} \cdot \vec{x}^{(s)})} \alpha_j^{(s)} \sum_{\lambda=1,2} [\epsilon_j(\lambda, k) \alpha(\lambda, k) \right. \\ \left. + \epsilon_j^*(\lambda, -k) \alpha^*(\lambda, -k)] dk^{(s)} \right\} \psi. \quad (26.24)$$

Often it is convenient to introduce discrete, instead of continuous, values for k .

If the term $H_1 \psi$ was not present in eq. (26.23), then we would be dealing with material particles which exert electrostatic forces on one another and a free radiation field, independent of these particles. The term $H_1 \psi$ thus determines the coupling of the particles with the radiation field and is decisive for a description of emission, absorption and dispersion of light. The success of non-relativistic quantum mechanics rests essentially on the assumption that the coupling with the radiation field can be considered as a relatively small perturbation. In the following we shall discuss how far the present theory corresponds to this requirement. It should be stressed that the equation (26.23) forms essentially the basis of the original Dirac theory of interaction between radiation and matter.⁴

Instead of using the $N(k)$ as the arguments of the ψ -function, we can introduce, following Landau and Peierls,⁵ the k -space of the photon. Then we have, for each of the total number N of photons present, the functions

$$f^{(M)}_{p_1 \dots p_m j_1 \dots j_N} (\vec{x}^{(1)}, \dots, \vec{x}^{(M)}, \vec{k}^{(1)}, \dots, \vec{k}^{(M)}).$$

The spin index p_i of the material particles runs from 1 to 4 and each index j for the photons runs from 1 to 3. The result of applying the Hamiltonian operator on these functions is obtained according to eq. (25.51) directly, if the form (26.22₁) of the operator $\alpha_j^{(s)}$ is used.

Applications: The application of the theory of interaction between radiation and matter based on the quantisation of the radiation field (i.e. the concept of the photon) rests on the understanding that in the coupling of radiation and matter in eq. (26.23) the perturbation H_1 is relatively small. Formally this perturbation calculation corresponds to a series expansion in powers of charge e .

As regards the results for the phenomena of emission, absorption and dispersion, we refer the reader to the relevant literature.* An essential success that this theory has achieved consists in the possibility of determining the radiation damping correctly (and hence the question of the line breadth).

* P.A.M. Dirac, Proc. Roy. Soc. London, 114, 243, 710 (1927). Dirac employed standing waves and not progressive waves for quantising the radiation. Secondly, at that time his theory of the electron had not appeared and, therefore, he used in the Hamiltonian for each particle the non-relativistic operator $(1/2m) \sum \vec{p}_j^2$ instead of $c (\sum \epsilon_j \vec{n}_j + m c \vec{B})$. This is approximately correct for small velocities of particles. The fact that the equations (26.23) and (26.24) follow from quantum electrodynamics was shown by J.R. Oppenheimer, Phys. Rev., 35, 461 (1930) and E. Fermi, Lincei Rend., 12, 431 (1930).

⁴ L. Landau and R. Peierls, Z. Phys. 62, 188 (1930).

⁵ G. Wentzel, Wave Mechanics of Collision and Radiation Processes [in German, Handbuch der Physik, edited by Geiger and Scheel, Vol. 24, Part I, Springer Verlag, Berlin (1933)]. See also W. Heitler, The Quantum Theory of Radiation, Third edition, Oxford University Press (1957).

We shall now show in brief how the apparently arbitrary Prescriptions I and II used in the semi-classical theory of the radiation (Sec. 15) can be justified by the photon theory. As regards the emission or scattering of radiation, it is not necessary to use the Hamiltonian operator directly, but we can, following Heisenberg,⁶ integrate the Maxwell equations, considered to be operator equations, by introducing retarded potentials. Here it is convenient to introduce instead of the electron position, the unperturbed stationary states of the system as variables and hence to decompose the function $\psi(\vec{X}_1 \dots \vec{X}_n, N(k, \lambda), t)$ in terms of the eigenfunctions $u_i(\vec{X}_1 \dots \vec{X}_n)$ of these states

$$\psi(\vec{X}_1 \dots \vec{X}_n, N(k, \lambda), t) = \sum_i c_i(N(k, \lambda), t) u_i(\vec{X}_1 \dots \vec{X}_n).$$

The calculation of the emitted or scattered radiation then follows exactly as in the semi-classical method; only the operators and the matrices, in general, act also on the photon variables.

Prescription I follows immediately from the fact that the expectation value of the intensity of radiation is given by $2F^*F$ and not by $F^*F + FF^*$. The decomposition of the field strengths into parts proportional to $e^{+i\omega t}$ and to $e^{-i\omega t}$ which was carried out in Sec. 15 [see eq. (15.13)], corresponds exactly to the decomposition of the field strengths according to eq. (25.09) into parts containing F^* only and parts containing F only. (This is true at least at such points in space where the charge and current densities vanish and this happens at large distances from the atom.) *The whole method underlying Prescription I amounts to the construction of the expectation value of F^*F in the photon theory.*

The basis for Prescription II which was required to calculate the reaction of the radiation on the atom is of a similar nature. We had found in eqs. (15.35) and (15.36) for the perturbed state amplitudes $c_m^{(1)}$ the expression:

$$c_m^{(1)} = \sum_n (\vec{f}_{mn}(t) \vec{E}^{(+)} + \vec{f}_{nm}^*(t) \vec{E}^{(-)}).$$

Instead of $E^{(+)}$ and $E^{(-)}$ we can also write $F^*(k)$ and $F(k)$. Then we have to remember that the $c_m^{(1)}$ are operators (or matrices) which act also on the photon variables. We have to calculate $\sum_{N'} |c_m^{(1)}(N', t)|^2$ which is equal to

$$\sum_{N'} |c_m^{(1)}(N', N)|^2$$

since $c_m^{(1)}(N', t) = \sum_N c_m^{(1)}(N', N) c_m^{(0)}(N, t)$, where N is the number of photons in the initial state. If we set

$$c_m^{(1)} = -i \left(\sum_n \vec{f}_{mn}^*(t) \vec{E}^{(-)} + \vec{f}_{nm}(t) \vec{E}^{(+)}, \right)$$

then

$$c_m^{(1)}(N, N') = (c_m^{(1)}(N', N))^*.$$

⁶ W. Heisenberg, Ann. d. Phys. (5), 9, 338 (1931). In contrast to Heisenberg's approach, we do not use here the method of quantisation of matter waves. The interaction between the electrons of the atom can then be arbitrary.

Hence

$$\sum_{N'} |c_m^{(1)}(N', N)|^2 = \sum_{N'} c_m^*(N, N') c_m(N', N) = (c_m^* c_m)_{N, N}$$

We have, therefore, to form the expectation value of $c_m^* c_m$ where c_m^* stands to the left of c_m . Then the expectation value of $2 \mathbf{F}^* \mathbf{F}$ is proportional to $|k| N(k)$ and that of $2 \mathbf{F} \mathbf{F}^*$ is proportional to $|k| \cdot N(k+1)$. The factor $2hv^3/c^3$ which was given earlier follows in a well-known way from the knowledge of the density of states of the eigen-vibrations.

By introducing the photon numbers as independent variables in the wave functions, the two special prescriptions of the semi-classical treatment of radiation processes based on the correspondence principle, are reduced to the general principles of wave mechanics.

As another problem which can be handled with the help of this theory, we can mention the Geiger-Bothe experiment on Compton effect which tests the correlation of the time points at which the scattered electron and the photon occur in the counters. From the theory it follows immediately that these time points must coincide within the limits of validity of geometrical optics and the limits obtained from the spreading of the wave-packets which define the initial state.⁷

27. Self-Energy of the Electron. Limits of the Present Theory

Even while calculating the electrostatic energy of particles from quantum electrodynamics we have seen that the electrostatic self-energy of individual particles turns out to be infinite. Further, it is observed that even if we set aside this infinitely large amount of energy, the resulting equations (26.22) and (26.24) still lead to an infinite magnetic self-energy.

If in the case of a single particle we make an expansion, in powers of the charge e of the particle, we obtain, according to the general formula of eq. (10.7b) for the perturbation calculation for a stationary state, an additional energy proportional to e^2 :

$$\Delta E = \sum_{n, k, \lambda} \frac{|H_{m0; n1\lambda}^{(1)}|^2}{E_m - E_n + \hbar c |k|}.$$

Here $H^{(1)}$ is the operator defined through (26.24). The indices m, n relate to the initial and intermediate states of the particle, and 0 and 1λ are the photon numbers in the initial and intermediate states. If we start from an initial state where no photon is present, then only those matrix elements, for which there is one photon in the intermediate state, are different from zero. Substitution of the operator $H^{(1)}$ in the above equation gives

⁷ Cf. W. Heisenberg, The Physical Principles of Quantum Mechanics, loc. cit.

$$\Delta E = e^2 \sum_n \int dk^{(3)} \frac{hc}{2|k|} \frac{1}{E_m - E_n + \hbar c |k|} \sum_{\lambda=1,2} \sum_{l=1}^3 \sum_{j=1}^3 u_m^*(\vec{x}) a_l u_n(\vec{x}) \\ \times u_n^*(x') a_j u_m(x') \vec{\epsilon}_l^* \cdot \vec{\epsilon}_j \cdot e^{i(\vec{k}(\vec{x}-\vec{x}'))} dx^{(3)} dx'^{(3)}.$$

In the free-particle case we can set

$$u_n(x) = a_n e^{i\vec{k}\vec{x}}$$

where ρ is the spin index and the role of the index n is taken up by K and the spin states (two positive and two negative energy states). If we are dealing with a bound particle, it turns out that the nature of the bound state does not matter if we are interested only in the way ΔE will become infinite. For this, in the free-particle case large values of K and in the general case the highly excited quantum states are important. There, however, in a region where $u_m(x)$ is noticeably different from zero, $u_n(x)$ can be replaced by the free-particle eigenfunction, if we substitute for K a mean value of the wave number of the n^{th} state in the region of the initial state m . It is, therefore, sufficient to consider the free-particle case.¹

We have now to sum, first, over λ and next over the four possible states of the material particle for given \vec{K} and integrate over \vec{x}' , \vec{x} and \vec{k} . The calculation gives, as Waller² and Oppenheimer³ have shown, an integrand which is dependent on \vec{k} and which becomes infinite for infinite \vec{k} . ΔE contains two terms. After integration over the directions of \vec{k} it has the form

$$\Delta E = f_1(|K_m|) \int_0^\infty |k| dk + f_2(K_m) \int_0^\infty dk.$$

Here f_1 is different from zero for an initially stationary particle, while f_2 is proportional to $|K_m|^2$ for small K_m and hence also proportional to the square of the velocity. The first (positive) term corresponds to the spin energy and the second (negative) term to that of magnetic field produced by the translatory motion.

There is no direct connection between the difficulty regarding an infinite self-energy and the state of negative energy discussed earlier. For even in theories that allow only positive energy (i.e. theories which exclude the intermediate state of negative energy like Schrödinger's and those which replace the Hamiltonian $\sum_k a_k \pi_k + mc \beta$ by $mc^2 + \frac{1}{2m} \sum_k \pi_k^2$) an infinite self-energy results.

It is further to be stressed that the difficulty of the infinite self-energy occurs in the same approximation (terms proportional to e^2) which is required in the theory of dispersion phenomena. We avoid this by restricting ourselves in the discussion of the time-dependent solution of the ψ -function to those terms which contain a resonance energy denominator and which grow fast with time.

The difficulty of self-energy arises essentially from the short waves of the radiation field and hence from small regions of space. We may ask whether a change in the

¹ I thank Prof. R. Peierls for these remarks. For the case of the self-energy of an oscillator, cf.; also L. Rosenfeld, Z. Phys., **70**, 454 (1931).

² J. Waller, Z. Phys., **62**, 673 (1930).

³ J.R. Oppenheimer, Phys. Rev., **35**, 461 (1930).

theory which corresponds to the introduction of a finite spread of the electron in classical electrodynamics, will help. This is indeed formally possible, but only on sacrificing the relativistic invariance of the theory.

Without changing the Hamiltonian (26.9) and the expression (26.10) for the momentum, we can introduce a "shape (Gestalt) function" $D(\vec{x})$ of the electron into the C.R. of $\pi_i^{[s]}$ with one another and with the electrical field strength \vec{E} . $D(\vec{x})$ is significantly different from zero only for distances of the order of the classical electron radius $d = \frac{c^2}{mc^2}$ (otherwise it is arbitrary).⁴ Then the C.R. (26.2) and (26.8) are to be replaced by

$$[\pi_j^{[s]}, \pi_i^{[s']}] = \delta_{ss'} - \frac{\hbar}{i} \frac{e}{c} \int D(\vec{x} - \vec{X}^{[s']}) H_{ij}(\vec{x}) dx^{[s]}, \quad (26.2')$$

$$[\pi_i^{[s]}, E_j(x)] = (-e) \frac{\hbar}{i} (-\delta_{ij}) \cdot D(\vec{x} - \vec{X}^{[s]}). \quad (26.8')$$

The subsidiary condition (26.7₂) is replaced by

$$\operatorname{div} \vec{E} = (-e) \sum_{s=1}^n D(\vec{x} - \vec{X}^{[s]})$$

and the equations of motion are

$$\frac{d\pi_k}{dt} = (-e) \int D(\vec{x} - \vec{X}^{[k]}) \left\{ E_k(\vec{x}) + \sum_l H_{kl}(\vec{x}) \alpha_l^{[s]} \right\} dx^{[s]}.$$

Finally, the Maxwell equation for the current is

$$-\frac{1}{c} \frac{\partial \vec{E}(\vec{x})}{\partial t} + \operatorname{rot} \vec{H}(\vec{x}) = (-e) \sum_{s=1}^n \vec{\alpha}^{[s]} D(\vec{x} - \vec{X}^{[s]}).$$

The self-energy will be finite since the D -function introduces a factor which practically makes the integrand in k -space to vanish for $k \gg 1/d$ (form factor of the electron).

Since, however, the relativistic invariance of the theory is lost as in the case of the corresponding method in classical theory – in the primed co-ordinate system the D -function would contain the time explicitly – it appears that this method is hardly acceptable as a solution of the difficulty.

The circumstance that the self-energy is infinite according to the theory prevents a subsequent relativistic treatment of the many-body problem.⁵

It is to be remarked that analogous to the case of the electron, an infinite self-energy occurs for the gravitational field, produced by a photon, if the field is quantised, though in the classical theory no point-singularity is introduced.⁶ This is

⁴ Cf. M. Born and G. Rumer, Z. Phys., **69**, 141 (1931).

⁵ Approximate ansatz for these have been presented in G. Breit, Phys. Rev., **34**, 553 (1929); **36**, 383 (1930) (magnetic interaction, but without retardation). See H. Bethe and E. Salpeter, Quantum Mechanics of One- and Two-electron problem, loc. cit. and C. Moller, Z. Phys., **70**, 786 (1931) (Collisions with weak interaction, treated as retarded), also Ann. d. Phys., **14**, 531 (1932). See further G. Wentzel, Wave Mechanics of Collision and Radiation Processes, loc. cit. Cf. also A.D. Fokker, Physica, **12**, 145 (1932) and Z. Phys., **58**, 386 (1929), where a two-body problem with partly retarded and partly advanced potentials is treated. For this problem, no radiation is emitted classically.

⁶ L. Rosenfeld, Z. Phys., **65**, 589 (1930). J. Solomon, ibid., **71**, 162 (1931).

connected with the fact that, as already stressed in Sec. 26, the use of the wave-mechanical formalism to a system with infinitely many degrees of freedom violates the correspondence principle. *Herein we may see a hint that not only the field concept but also the space-time concept in small regions may require a fundamental modification.*

We may make the following remarks on the necessary modifications of the field concept: The present theory rests on two logically independent bases to which in classical theory correspond point mechanics and the Maxwell theory. These bases are the wave mechanics of the material particle and the wave mechanics of the cavity radiation (photon theory). To this corresponds the fact that the present theory does not take the atomic nature of the electric charge into account, since it could be brought into agreement with arbitrarily many and also arbitrarily small electrical charges. To this also corresponds the (unsuccessful) attempt of the present theory to resolve the whole eigen-field of a moving electron into the photons instead of interpreting it as forming an indivisible whole connected with the electron. This field is inseparably associated with a definite numerical value of a dimensionless number, viz., the fine structure constant $e^2/\hbar c$. A future complete relativistic quantum theory must bring about a deep unification of the foundations.*

* After the early developments of Quantum Electrodynamics in the hands of Dirac, Pauli, Heisenberg, Jordan and Wigner (as presented in this book), there was a period of frustration and waiting before the theory could proceed further. In 1947 Kramers suggested a remedy to some of the evils, which the theory is heir to, by introducing the notion of mass renormalisation, viz., identifying the measured mass of the electron with that of the electron interacting with its radiation field and not with that of the electron decoupled from the electromagnetic field. This stimulated Bethe to compute correctly (but using a non-relativistic procedure) the Lamb shift, viz., the shift between the $2S^{1/2}$ and $2P^{1/2}$ levels of the hydrogen atom which in the simple Dirac theory are degenerate. Relativistic calculations were carried out by Schwinger and Tomonaga with the help of field theory and by Feynman utilising the notion of particles interacting with themselves and with one another through a propagated action at a distance (leading to the very fruitful concept of Feynman diagrams). The equivalence of the two approaches was soon proved by Dyson. With these advances all the basic electrodynamic processes like Compton effect and electron-electron scattering, the Lamb shift, positronium, the anomalous magnetic moment of the electron and more complicated phenomena such as the Lamb shift of the hyperfine structure and the exact shape of the spectral lines could be explained after the mass and charge renormalisations had been accomplished. The agreement with experiment was remarkable. In principle there should be no further infinities in the theory. Källén has given a proof, using a non-perturbative approach, that at least one of the renormalisation constants has to be necessarily infinite. However, his proof has been criticised as being gauge-dependent.

Landau [L.D. Landau in "Niels Bohr and the Development of Physics", Pergamon Press, London (1955)] has pointed out further inconsistencies which can arise in the theory, but these are expected to be relevant only at very high energies.

The epoch-making papers of the crucial periods in the progress of Quantum Electrodynamics have been collected by Schwinger [Julian Schwinger (ed.): Selected Papers on Quantum Electrodynamics, Dover Publications, Inc., New York (1958)] along with an illuminating preface.

Some further references are:

S.S. Schweber, H.A. Bethe and F. de Hoffmann: Mesons and Fields, Vol. I, Row, Peterson & Co., Evanston, Illinois (1953); J.M. Jauch and F. Rohrlich, Theory of Photons and Electrons, Second Expanded Edn., Springer Verlag, New York (1976); J.D. Bjorken and S.D. Drell, Relativistic Quantum Fields, McGraw-Hill Book Co., New York (1965); G. Källén, Quantum Electrodynamics (translated from German by C.K. Iddings and M. Mizushima) George Allen & Unwin Ltd., London and Springer Verlag, New York (1972). The last book contains a lucid exposition of the subject in the Heisenberg representation.

Bibliography

The number of books on Quantum Mechanics presently available is quite large. The following list does not aim at providing an exhaustive bibliography, but only gives a short selection (Part A), in addition to the articles and books already referred to in the text - which, for convenience are collected under Part B below, all the references being arranged in the alphabetical order of the authors.

Part A

- Baym, Gordon: *Lectures on Quantum Mechanics*, W.A. Benjamin Inc., Reading, Massachusetts (1973).
- Bethe, H.A.: *Intermediate Quantum Mechanics*, 2nd edn. W.A. Benjamin Inc., New York (1966).
- Biedenharn, L.C. and Van Dam, H. (Eds.): *Quantum Theory of Angular Momentum (A Collection of Reprints and Original Papers)*, Academic Press (1965).
- Bjorken, J.D. and Drell, S.D.: *Relativistic Quantum Mechanics*, McGraw-Hill Book Co., Toronto (1964).
- Bohm, David: *Quantum Theory*, Asia Publishing House, Bombay (1962).
- Condon, E.U. and Shortley, G.H.: *The Theory of Atomic Spectra*, University Press, Cambridge (1953).
- Feynman, R.P.: *Quantum Electrodynamics*, W.A. Benjamin Inc., New York (1962).
- Feynman, R.P. and Hibbs, A.R.: *Quantum Mechanics and Path Integrals*, McGraw-Hill Book Co., New York (1965).
- Feynman, R.P.: *The Theory of Fundamental Processes*, W.A. Benjamin Inc., New York (1961).
- Gottfried, K.: *Quantum Mechanics*, W.A. Benjamin Inc., New York (1966).
- Jammer, M.: *The Philosophy of Quantum Mechanics*, Wiley-Interscience Publn., New York (1974).
- Jauch, Joseph M.: *Foundations of Quantum Mechanics*, Addison-Wesley, Mass., U.S.A. (1968).
- Kramers, H.A.: *Quantum Mechanics* (translated by D. ter Haar), North-Holland Pub. Co., Amsterdam (1958).
- Kursunoğlu, Behram: *Modern Quantum Theory*, W.H. Freeman and Co., San Francisco (1962).
- Lipkin, Harry J.: *Quantum Mechanics*, North-Holland Pub. Co., Amsterdam; American Elsevier Pub. Co., New York (1973).
- Mackey, G.W.: *The Mathematical Foundations of Quantum Mechanics*, W.A. Benjamin, Inc., Amsterdam (1963).
- Mandl, F.: *Quantum Mechanics*, Butterworths Scientific Publications, London (1958).
- March, A.: *Quantum Mechanics and Wave Fields*, John Wiley & Sons, New York (1951).
- Mathews, P.M. and Venkatesan, K.: *A Text Book of Quantum Mechanics*, Tata McGraw-Hill Pub. Co. Ltd., New Delhi (1977).
- Messiah, A.: *Quantum Mechanics*, Vol. I, II, translated by J. Potter, North-Holland Publishing Co., Amsterdam (1964).
- Mott, N.F. and Massey, H.S.W.: *The Theory of Atomic Collision*, 3rd edn., Clarendon Press, Oxford (1965).
- Mott, N.F. and Sneddon, N.I.: *Wave Mechanics and its Applications*, Clarendon Press, Oxford (1948).
- Pauli, W.: *Wave Mechanics (Pauli Lectures on Physics Series, Vol. 5 - ed. C.P. Enz)* translated by H.R. Lewis and S. Margulies, The MIT Press, Cambridge (1973).

- Pauling, L.C. and Wilson, B.E.: *Introduction to Quantum Mechanics with Applications to Chemistry*, McGraw-Hill Book Co., New York (1935).
- Petraschen, M.I. and Trifonov, E.D.: *Application of Group Theory in Quantum Mechanics*, translated by S. Chomet, Iliffe Books Ltd., London (1969).
- Powell, J.L. and Craemann, B.: *Quantum Mechanics*, Addison-Wesley Pub. Co., Inc., Reading, Mass. U.S.A. (1961).
- Prugovecki, E.: *Quantum Mechanics in Hilbert Space*, Academic Press, New York (1971).
- Sakurai, J.J.: *Advanced Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1967).
- Theo Kahan: *Theory of Groups in Classical and Quantum Physics*, translated by H. Ingram, Oliver and Boyd, Edinburgh (1965).
- Trigg, G.L.: *Quantum Mechanics*, D. Von Nostrand Co., Inc., Princeton, N.J. (1968).

Part B

- Abdus Salam and Wigner, E.P. (Eds.): *Aspects of Quantum Mechanics*, Cambridge University Press, Cambridge (1972).
- Akhiezer, A.I. and Berestetskii, V.B.: *Quantum Electrodynamics*, Interscience Publishers, New York (1965).
- Band, W.: *An Introduction to Quantum Statistics*, D. Von Nostrand Co., Inc., Princeton, New Jersey (1955).
- Bates, D.R. (Ed.): *Quantum Theory*, Vol. I, Academic Press, New York (1962).
- Bethe, H.A. and Salpeter, E.E.: *Quantum Mechanics of One- and Two-Electron Atoms*, Springer Verlag, Berlin (1957).
- Bjorken, J.D. and Drell, S.D.: *Relativistic Quantum Fields*, McGraw-Hill Book Co., New York (1965).
- Blokhintsev, D.I.: *Quantum Mechanics*, (translated from Russian, 4th edn. by Sykes, J.B. and Kearsley, M.J.), Reidel Pub. Co, Dordrecht, Holland (1964).
- Bohr, N.: *Atomic Theory and the Description of Nature*, Cambridge University Press, Cambridge (1961).
- Brillouin, L.: *Die Quantenstatistik*, Berlin (1931).
- Courant, R. and Hilbert, D.: *Methods of Mathematical Physics*, Vol. I, Interscience Publishers, Inc., New York (1975).
- de Broglie, L.: *Introduction à l'étude de la mécanique ondulatoire*, Paris (1930).
- Dirac, P.A.M.: *The Principles of Quantum Mechanics*, 4th edn., Oxford University Press, Oxford (1959).
- Edmonds, A.R.: *Angular Momentum in Quantum Mechanics*, Princeton University Press, Princeton, New Jersey (1960).
- Flügge, S.: *Practical Quantum Mechanics*, Vol. I, Springer Verlag, Berlin (1971).
- Green, H.S.: *Matrix Mechanics*, P. Noordhoff Ltd., Netherlands (1965).
- Hamermesh, M.: *Group Theory and Its Applications to Physical Problems*, Addison-Wesley, Reading, Mass., U.S.A. (1962).
- Heisenberg, W.: *The Physical Principles of the Quantum Theory*, Dover Publications, Inc., New York (1949).
- Heitler, W.: *The Quantum Theory of Radiation*, 3rd edn., Oxford University Press, Oxford (1957).
- Jauch, J.M. and Rohrlich, F.: *Theory of Photons and Electrons*, Second Enlarged Edn., Springer Verlag, New York (1976).
- Jordan, P.: *Statistische Mechanik auf quantentheoretischer Grundlage*, Braunschweig (1933).
- Källén, G.: *Quantum Electrodynamics*, (translated from German by C.K. Iddings and M. Mizushima) George Allen & Unwin Ltd., London and Springer Verlag, New York (1972).
- Landau, L.D.: in, "Niels Bohr and the Development of Physics", Pergamon Press, London (1955).
- Landau, L.D. and Lifschitz, E.M.: *Quantum Mechanics*, Pergamon Press, London (1958).
- Larmor: *Aether and Matter*, Cambridge (1900).
- Lighthill, M.J.: *Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press, Cambridge (1964).

- Max Born and Jordan, P.: Elementare Quantenmechanik, Springer Verlag, Berlin (1930).
- Merzbacher, E.: Quantum Mechanics, 2nd edn., John Wiley & Sons, Inc., New York (1970).
- Neumann, J.v.: Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton (1955).
- Particle Data Group: Review of Particle Properties, Review of Modern Physics, April (1978).
- Pauli, W.: Continuous Groups in Quantum Mechanics, Cern Publications (1956), Springer Tracts in Modern Physics 37, Berlin (1965).
- Pauli, W.: in "Niels Bohr and the Development of Physics", Pergamon Press, London (1955).
- Racah, G.: Group Theory and Spectroscopy in Springer Tracts in Modern Physics 37, Berlin (1965).
- Schiff, L.I.: Quantum Mechanics, 3rd edn., McGraw-Hill Book Co., New York (1968).
- Schrödinger, E.: Lectures on Wave Mechanics (in German), Berlin (1928).
- Schweber, S.S., Bethe, H.A. and de Hoffmann, F.: Meson and Fields, Vol. I, Row, Peterson & Co., Evanston, Illinois (1955).
- Schwinger, J.: (Ed.): Selected Papers on Quantum Electrodynamics, Dover Publications, Inc., New York (1958).
- Sommerfeld, A.: Wave Mechanics, Dutton, New York (1929).
- Streater, R.F. and Wightman, A.S.: "PCT, Spin and Statistics, and All That", W.A. Benjamin, New York (1964).
- Van der Waerden, B.L.: Die Gruppentheoretische Methode in der Quantenmechanik, Springer Verlag, Berlin (1932).
- Wentzel, G.: Wave Mechanics of Collision and Radiation Process (in German) Handbuch der Physik, edited by Geiger and Scheel, Vol. 24, Part I, Springer Verlag, Berlin (1933).
- Weyl, H.: Group Theory and Quantum Mechanics, 2nd edn., Leipzig (1931); Dover Publications Inc., New York (1950).
- Wigner, E.P.: Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra, Academic Press, New York (1959).
- Wintner, A.: Spektraltheorie der unendlichen Matrizen, Leipzig (1929).

• • •

• • •

Index

- Adiabatic principle, Ehrenfest's 86
— perturbation 86
Alpha particle 117, 125
Angular frequency 3
Angular momentum 46, 47, 107
— commutation relations 107
— conservation law 191
— conservation law in Dirac theory 167
— matrix elements 108
— operator 46, 47, 107
— operators — application to 190
Anticommutation relation 115, 149, 150
Antisymmetric tensor 195
— wave function 118
Approximation methods 79
- Bohr's correspondence principle 25, 100
Bose-Einstein Statistics 125
Bosons 115
- Charge renormalisation 204
Classical mechanics as a limiting case 172
Commutation relations
— angular momentum 107
— energy 63
— field strength vectors 183, 185
— position and momentum operators 37, 181
Complementarity 1, 7
Complex field strength vector 177
Compton effect 201
Configuration space of photons, transition to 189
Continuity equation 16, 25, 93
— of the Dirac theory 147
Continuous group 106
Correspondence principle 25, 100
Cut off in k-space 187
- Darwin term 169
de Broglie wavelength 6, 10
- de Broglie waves 4, 6
de Broglie's hypothesis 4
Degenerate eigenfunction 49
Degeneracy, degree of 48
Degeneracy, removal by perturbation 82
Density matrix 72
Deuteron 116
Diffraction experiment 1
— grating 1
— of corpuscular rays 4
Dipole radiation 136
Dirac delta function 52, 193
— equation with external electromagnetic field 163
— equation for the free electron 148
— matrices 148, 149, 150
Dirac theory, Approximations to 168
—, Limitations of 174
Dirac's theory of radiation 130
— the electron 145
Discrete eigenvalues 43
Dispersion 136
Doppler effect 11
- Ehrenfest's adiabatic law 86
Eigenfunctions 44, 48
Eigenfunction of identical particles, symmetry of 116
Eigenvalue problem 42
— spectrum 51
Eigenvibration in a cubical cavity 177
Electromagnetic waves 3
Electron 130
— Accuracy of field measurement with 188
—, Dirac's theory of 145
—, Self energy of 201
Electric and magnetic field strength vectors 177
Elementary particle 116
Emission of radiation 134
Energy exchange between radiation and matter 130
Energy of progressive waves 178
Equivalence of operators and matrices 55
Exchange integral 123

- Exclusion principle 116, 119, 120
 Expectation values 38, 56, 72
 — of perturbation operator 80
 External force field 25
 — magnetic field 28
 Euler angles 48, 111

 Fermi-Dirac Statistics 125
 Fermions 115
 Fermi, method of 196
 Field strengths, Mean of 186
 Field strengths, limits on the accuracy of measurement of 187
 Fourier-integral for wave functions of free particles 14
 Free electron, Dirac theory of 148
 Free radiation field, quantisation 177

 Galilean transformation 19
 Gamma ray microscope 72
 Gauge invariance 31
 — Group 31
 — transformation 30
 Geiger-Bothe experiment on Compton Effect 201
 Geometrical optics, region of validity 191
 Group, continuous 105
 — of gauge transformation 196
 — of unitary transformations of two complex variables 183
 —, Lorentz 107
 —, Rotation 106, 108
 —, theory 103
 —, transformation 103
 —, velocity 5, 160

 Hadrons 116
 Hamiltonian for a system of point masses 25
 — for a particle in a magnetic field 28
 — of the Dirac theory 164
 — functions 103
 — operator 34
 — in spherical polar co-ordinates 41
 Hamilton-Jacobi partial differential equation 92, 173
 Heisenberg equations of motion 61
 Heisenberg's theory of the helium spectrum 123
 — uncertainty principle 1, 4
 Helium 116
 Helium spectrum, Heisenberg's theory 123
 Hermitian matrix 54
 Hermiticity 24, 25
 Hole theory 176
 Hydrogen atom, Schrödinger equation for 46
 — in Matrix Mechanics 59

 Identical particles 32, 116
 Improper function 52, 184, 185
 Induced emission 140
 Infinite degrees of freedom, system with 186
 Infinitesimal transformation group 105
 — transformations of first and second kind 195
 Interference experiments 2
 Invariance of the Hamiltonian 103
 Irreducible representation 105, 183

 Jerk effect 87
 Jittery motion 160
 Jordan-Wigner anticommutation relations 127

 Klein-Gordon equation 146
 Klein-Nishina formula 166
 Klein's paradox 174

 Landau and Peierls, argument of 187
 Laplacian operator 14
 Laporte's rule 112
 Laws of motion, general forms of 61
 Lepton 116
 Light emission 134
 Linearity of an operator 15
 Lorentz group 107
 — invariance 150
 — transformation, behaviour under 186

 Many-particle interactions 31
 Mass renormalisation 204
 Matrix mechanics 54
 — transformation 54
 —, unitary 57
 Matter waves 4, 6
 Maxwell equation, integration using retarded potentials 200
 —, solution of 178
 — when charges are present 197
 Measurements, concept of 67
 —, theory of 67
 — of the first kind 75
 — of the second kind 75
 Method of Fermi 196
 Molecular beam experiment 68
 Momentum measurement 8, 75
 — of progressive waves 182
 Momentum-space integral 114
 Multipole radiation 135

 Needle radiation 143
 Negative energy levels 176
 —, transitions to 174

- Neutrino 157
 Non-relativistic wave equation for the many-body problem 35
 Normalisation of eigenfunctions 49
 Nucleon 116
 Observable 56
 Old quantum theory, relation to 98
 One-electron atom 45
 One-particle problem, relativistic 145
 Operator 15
 — calculus 31
 —, creation, annihilation 181
 —, Hamiltonian 34
 —, Linear 15, 34
 —, Number 184
 Operator, unitary 57
 Operators and matrices, equivalence of 55
 Orbital angular momentum 112, 168, 174
 Orthogonality of eigenfunctions 48
 Oscillator, harmonic 45
 —, isotropic 45
 Parity 112
 Particles, elementary 116
 —, identical 32, 116
 Pauli approximation, non-relativistic 168
 — exclusion principle 116, 119, 120
 — spin matrices 115, 149
 Permutation 116
 Perturbation theory, adiabatic 86
 —, energy, expectation value 80
 —, sudden 86
 — theory 79
 —, stationary state 79
 —, time-dependent 83
 Phase integral 97
 —, method of quantisation 180
 Photon 1, 2, 3, 4, 5, 143
 Photons, no current density for 191
 —, number of 183
 —, transition to the configuration space of 187
 Planck's constant 3
 Polarisation of electron waves 174
 — of photon 177
 Polarised eigen-vibrations 179
 Position measurement 8
 Positron 176
 Potential barrier 98
 Probability 10, 12, 15, 16
 — current 17, 25
 — density in co-ordinate space 10, 15, 24
 — density in momentum space 12, 15
 —, general statement in quantum mechanics 89
 Progressive waves 177, 180
 Projection operator 52
 Proton 32, 112, 116
 Quadrupole radiation 136
 Quantisation of charge 204
 Quantisation of radiation field 180
 Quantum conditions 97
 — electrodynamics 177
 —, number 52, 98, 100, 101
 Quantum theory, relativistic 145
 Radiation 130, 166, 199
 —, coherence properties of 140
 —, damping 142, 144
 —, emission of 134
 — field considered as small perturbation 199
 —, linearly polarised 179
 —, scattering of 137, 166
 Radiation processes, treatment using correspondence principle 130
 — semiclassical theory 130, 201
 Radiation and matter, application of 199
 —, interaction between 192
 Recoil through light emission 142, 143
 Reducible representation 104
 Reflection invariance 154
 Relativistic invariance of the Dirac Hamiltonian 150
 — of the theory 194
 Representation of a transformation group 104
 Rotation group, three dimensional 106, 108
 Russel-Saunders coupling 121
 Rydberg constant 45
 Scattering of radiation 137, 166
 Schrödinger equation for the hydrogen atom 45
 —, time dependent 24
 Schrödinger's wave equation 24
 Schrödinger wave function 23, 197
 Schrödinger, variational principle of 53
 Self energy of the electron 201
 Second quantisation 127
 Secular determinant 81
 Sélection rules for radiative transitions 135
 Semiclassical theory of radiation 130, 201
 Singlet states 122, 123
 Singularities 44
 Singular functions 179
 Slater determinant 120
 Space co-ordinates 121
 Space-time concept in small regions 204
 Spherical harmonics 46
 Spherical top 48
 Spin 103, 112, 165, 168, 170

- Spin co-ordinates 121
 — matrices 115, 149
 — moment 112, 168, 170
 Spinor 156
 — calculus 156
 Spontaneous emission 140
 Spur of a matrix 73
 Stationary states as eigenvalue problem 42
 —, determination through measurement 67
 Statistical application of quantum theory 124
 — Weight of a state 71
 — interpretation of wave function 8, 25
 Stern-Gerlach experiment 68
 Stieltjes integral 44
 Stress tensor 26, 36
 Subsidiary condition 193
 Sub-space with a finite number of co-ordinates 59, 81
 Sudden change of parameter 86, 89, 90
 Superposition principle 2, 24
 Symmetric eigenfunctions 118, 120
 — states 125, 189
 Thomas correction 169
 Three-dimensional rotation group 106, 108
 Time as a c-number 63
 — derivative of an operator 62
 Trace of a matrix 73
 Transformation group 103
 —, infinitesimal 106
 — of matrices 54
 — of operators 54
 Transition probability 84
 — process 72
 Transmission coefficient 175
 Transversality condition 178, 190
 Triplet states 123
 Uncertainty principle of Heisenberg 1, 4
 —, quantitative formulation of 21
 Uncertainty in momentum, additional 187
 Unitary matrix 57
 Unitary transformation 58
 — of two-complex variables 183
 Variational principle of Schrödinger 53
 Vector model 110
 Virial theorem 27
 Wave concept of light 1
 Wave equation, Schrödinger's 23
 —, relativistic, of Dirac 148, 163
 — for several particles, non-relativistic 35
 — in curvilinear co-ordinates 40
 — of second order 148
 — when forces are present 163
 Wave function, antisymmetric 118
 —, symmetric 118, 120
 — of a particle acted on by external forces 23
 — of a free particle 13
 Wave fields, classical 2
 Wave number vector 3
 Wave-packet 2, 3, 18, 101
 — behaviour, in free particle case 160
 —, reduction of 71
 Wave-particle duality 1
 W K B approximation for Dirac equation 172
 — for Schrödinger equation 91, 92
 Zero-point energy, avoiding the 185, 186
 —, infinite 179
 Zitterbewegung 160

J. M. Blatt, V. F. Weisskopf

Theoretical Nuclear Physics

1979. 126 figures, 36 tables. XIV, 864 pages
ISBN 3-540-90382-8

Contents: General Properties of the Nucleus. – Two-Body Problems at Low Energies. – Nuclear Forces. – Two-Body Problems at High Energies. – Three- and Four-Body Problems. – Nuclear Spectroscopy I: General Theory. II: Special Models. – Nuclear Reactions: General Theory. – Nuclear Reactions: Application of the Theory to Experiments. – Formal Theory of Nuclear Reactions. – Spontaneous Decay of Nuclei. – Interaction of Nuclei with Electromagnetic Radiation. – Beta-Decay. – Nuclear Shell Structure. – Appendices.

This classic is still considered to be one of the finest references in the field, despite the developments of the last 25 years. One of its most important features is its accessibility to the experimental physicist. The authors develop theoretical nuclear physics as a "semi-empirical" subject, a point of view still valid and appropriate today. They basically confine themselves to phenomena involving energies below about 50 MeV; nuclear phenomena in cosmic rays, as well as the phenomena associated with the production and absorption of mesons excluded.



Springer-Verlag
Berlin
Heidelberg
New York

Wolfgang Pauli

Wissenschaftlicher Briefwechsel mit Bohr,
Einstein, Heisenberg u. a. – Band I: 1919–1929
Scientific Correspondence with Bohr, Einstein,
Heisenberg a. o. – Volume I: 1919–1929

Herausgeber/Editors: A. Hermann, K. v. Meyenn,
V. F. Weisskopf
Vorwort von/Preface by V. F. Weisskopf

1979. 13 Abbildungen. Etwa 620 Seiten (davon
etwa 120 Seiten in Dänisch, etwa 10 Seiten in
Englisch).

(Sources in the History of Mathematics and
Physical Sciences, Volume 2)

ISBN 3-540-08962-4

Inhaltsübersicht: Das Jahr 1919: Auseinandersetzung mit der Allgemeinen Relativitätstheorie. – Das Jahr 1920: "Relativitätsartikel" und erste Arbeiten zur Atomphysik. – Das Jahr 1921: Dissertation über das Wasserstoffmolekül. – Das Jahr 1922: Göttingen – Hamburg – Kopenhagen. – Das Jahr 1923: Anomaler Zeeman-Effekt. – Das Jahr 1924: Weg zum Ausschließungsprinzip. – Das Jahr 1925: "Quantenartikel" und Göttinger Matrizenmechanik. – Das Jahr 1926: Rotierendes Elektron und Verallgemeinerungen der Quantenmechanik. – Das Jahr 1927: Kopenhagener Interpretation und Quantenelektrodynamik. – Das Jahr 1928: Berufung nach Zürich – Schwierigkeiten in der Quantenelektrodynamik. – Das Jahr 1929: Systematischer Aufbau der Quantenfeldtheorie. – Anhang.

Wolfgang Pauli (1900–1958) war einer der bedeutendsten Physiker des 20ten Jahrhunderts. Im Jahre 1945 wurde ihm der Nobelpreis für sein Werk verliehen. Besonders bekannt sind Pauli's Beiträge in der Quantentheorie, insbesondere das nach ihm benannte Ausschließungsprinzip.

Pauli's Korrespondenz ist von ausserordentlichem physikalischen und historischem Wert. Viele seiner Ideen und kritischen Beiträge wurden erstmals in Briefen entwickelt. Da Pauli mit allen führenden Begründern der Quantentheorie in regem Kontakt stand, vermitteln die Briefe ein lebendiges Bild der damaligen Problemstellungen und der Diskussion, die schließlich zu unseren heutigen Auffassungen geführt haben. In einem hohen Maße fungierte dabei Pauli als das unbestechliche Gewissen, das die Hypothesen und Theorien der Kollegen einer strengen Kritik unterzog. Darüberhinaus gehen auf Pauli viele originelle schöpferische Beiträge zurück, die auf Grund des Briefwechsels nunmehr erstmals im Zusammenhang verstanden werden können. Die vorliegende Ausgabe enthält nicht nur Pauli's Briefe, sondern auch, soweit vorhanden, diejenigen seiner Korrespondenten. Ein umfangreicher Kommentar erleichtert das Verständnis der Korrespondenz.

Texts and Monographs in Physics

Editors: W. Beiglböck, M. Goldhaber,
E. H. Lieb, W. Thirring

A. Böhm

Quantum Mechanics

1979. 105 figures, 7 tables. XVII, 522 pages
ISBN 3-540-08832-3

O. Bratteli, D. W. Robinson

Operator Algebras and Quantum Statistical Mechanics

Volume I

The Mathematical Theory of C^* - and W^* -algebras

1979. XII, 500 pages

ISBN 3-540-09187-4

H. M. Piltz

Relativistic Particle Physics

1979. 85 figures. XII, 427 pages

ISBN 3-540-09348-6

R. D. Richtmyer

Principles of Advanced Mathematical Physics I

1978. 45 figures. XV, 422 pages

ISBN 3-540-08873-3

R. M. Santilli

Foundations of Theoretical Mechanics I:

The Inverse Problem in Newtonian Mechanics

1978. 5 figures. IX, 266 pages

ISBN 3-540-08874-1

M. D. Scadron

Advanced Quantum Theory and Its Applications Through Feynman Diagrams

1979. 78 figures, 1 table. XIV, 386 pages

ISBN 3-540-09045-2

J. Kessler

Polarized Electrons

1976. 104 figures. IX, 223 pages
ISBN 3-540-07678-6

W. Rindler

Essential Relativity

Special, General, and Cosmological
Second Edition 1977. 44 figures
XV, 284 pages

ISBN 3-540-07970-X

K. Chadan, P. C. Sabatier

Inverse Problems in Quantum Scattering Theory

With a Foreword by R. G. Newton

1977. 24 figures. XXII, 344 pages

ISBN 3-540-08092-9

J. M. Jauch, F. Rohrlich

The Theory of Photons and Electrons

The Relativistic Quantum Field Theory
of Charged Particles with Spin One-Half
Second Expanded Edition 1976.

55 figures, 10 tables. XIX, 553 pages
ISBN 3-540-07295-0

C. Truesdell, S. Bharatha

The Concepts and Logic of Classical Thermodynamics as a Theory of Heat Engines

Rigorously Constructed upon the
Foundation

Laid by S. Carnot and F. Reech

1977. 15 figures. XXII, 154 pages

ISBN 3-540-07971-8



Springer-Verlag
Berlin
Heidelberg
New York

Pauli General Principles of Quantum Mechanics

Only about seven years after quantum mechanics obtained its final form W. Pauli, himself one of the most active and most critical pioneers of the new theory, wrote this classic. He revised it in 1958 for the Encyclopedia of Physics and it was this version from which this translation was made. More than any other book on quantum mechanics this text works out its foundations, which explains the continuing interest of the physics community in this book. For their historical value the translators included also the sections on quantum electrodynamics from the 1933 edition.

ISBN 3-540-09842-9
ISBN 0-387-09842-9