QUANTUM MECHANICS

L.D. LANDAU, E.M. LIFSHITZ

Contents

FROM THE PREFACE TO THE FIRST ENGLISH EDITION

THE present book is one of the series on Theoretical Physics, in which we endeavour to give an up-to-date account of various departments of that science. The complete series will contain the following nine volumes:

1. *Mechanics.*

2. *The classical theory of fields.*

3. *Quantum mechanics (non-relativistic theory).*

4. *Relativistic quantum theory.*

5. *Statistical physics.*

6. *Fluid mechanics.*

7. *Theory of elasticity.*

8. *Electrodynamics of continuous media.*

9. *Physical kinetics.*

Of these, volumes 4 and 9 remain to be written.

The scope of modern theoretical physics is very wide, and we have, of course, made no attempt to discuss in these books all that is now included in the subject. One of the principles which guided our choice of material was not to deal with those topics which could not properly be expounded without at the same time giving a detailed account of the existing experimental results. For this reason the greater part of nuclear physics, for example, lies outside the scope of these books. Another principle of selection was not to discuss very complicated applications of the theory. Both these criteria are, of course, to some extent subjective.

We have tried to deal as fully as possible with those topics that are included. For this reason we do not, as a rule, give references to the original papers, but simply name their authors. We give bibliographical references only to work which contains matters not fully expounded by us, which by

their complexity lie "on the borderline" as regards selection or rejection. We have tried also to indicate sources of material which might be of use for reference. Even with these limitations, however, the bibliography given makes no pretence of being exhaustive.

We attempt to discuss general topics in such a way that the physical significance of the theory is exhibited as clearly as possible, and then to build up the mathematical formalism. In doing so, we do not aim at "mathematical rigour" of exposition, which in theoretical physics often amounts to self-deception.

The present volume is devoted to non-relativistic quantum mechanics. By "relativistic theory" we here mean, in the widest sense, the theory of all quantum phenomena which significantly depend on the velocity of light. The volume on this subject (volume 4) will therefore contain not only Dirac's relativistic theory and what is now known as quantum electrodynamics, but also the whole of the quantum theory of radiation.

L.D. LANDAU and E.M. LIFSHITZ, Institute of Physical Problems, USSR Academy of Sciences
August 1956

PREFACE TO THE SECOND ENGLISH EDITION

FOR this second edition the book has been considerably revised and enlarged, but the general plan and style remain as before. Every chapter has been revised. In particular, extensive changes have been made in the sections dealing with the theory of the addition of angular momenta and with collision theory. A new chapter on nuclear structure has been added; in accordance with the general plan of the course, the subjects in question are discussed only to the extent that is proper without an accompanying detailed analysis of the experimental results.

We should like to express our thanks to all our many colleagues whose comments have been utilized in the revision of the book. Numerous comments were received from V. L. Ginzburg and Ya. A. Smorodinskii. We are especially grateful to L. P. Pitaevskii for the great help which he has given in checking the formulae and the problems.

Our sincere thanks are due to Dr. Sykes and Dr. Bell, who not only translated excellently both the first and the second edition of the book, but also made a number of useful comments and assisted in the detection of various misprints in the first edition.

Finally, we are grateful to the Pergamon Press, which always acceded to our requests during the production of the book.

L.D. LANDAU and E.M. LIFSHITZ
October 1964

PREFACE TO THE THIRD RUSSIAN EDITION

THE previous edition of this volume was the last book on which I worked together with my teacher L. D. Landau. The revision and expansion that we then carried out was very considerable, and affected every chapter.

For the third edition, naturally, much less revision was needed. Nevertheless, a fair amount of new material has been added, including some more problems, and relating both to recent research and to earlier results that have now become of greater significance.

Landau's astonishing grasp of theoretical physics often enabled him to dispense with any consultation of original papers: he was able to derive results by methods of his own choice. This may have been the reason why our book did not contain certain necessary references to other authors. In the present edition, I have tried to supply them as far as possible. I have also added references to the work of Landau himself where we describe results or methods that are due to him personally and have not been published elsewhere.

As when dealing with the revision of other volumes in the Course of Theoretical Physics, I have had the assistance of numerous colleagues who informed me either of deficiencies in the treatment given previously, or of new material that should be added. Many useful suggestions incorporated in this book have come from A. M. Brodskii, G. F. Drukarev, I. G. Kaplan, V. P. Krainov, I. B. Levinson, P. E. Nemirovskii, V. L. Pokrovskii, I. I. Sobel'man, and I. S. Shapiro. My sincere thanks are due to all of these.

The whole of the work on revising this volume has been done in close collaboration with L. P. Pitaevskii. In him I have had the good fortune to find a fellow-worker who has passed likewise through the school of Landau and is inspired by the same ideals in the service of science.

E.M. LIFSHITZ, Moscow
November 1973

EDITOR'S PREFACE TO THE FOURTH RUSSIAN EDITION

IN this edition of Quantum Mechanics some misprints and errors noted since the publication of the third edition have been corrected. Some small improvements have also been made, and several problems have been added.

I am grateful to all readers who have provided me with comments.

L.P. PITAEVSKI, May 1988

NOTATION

- Operators are denoted by a circumflex: $\widehat{f}$

- dV volume element in coordinate space

- dq element in configuration space

- d^3p element in momentum space

- $f_{nm} = f_{mn} = \langle n|f|m \rangle$ matrix elements of the quantity f (see definition in §11)

- $\omega_{nm} = (E_n E_m)/\hbar$ transition frequency

- $\left\{ \widehat{f}, \widehat{g} \right\} = \widehat{f}\widehat{g} - \widehat{g}\widehat{f}$ commutator of two operators

- $\widehat{H}$ Hamiltonian

- δ_l phase shifts of wave functions

- Atomic and Coulomb units (see beginning of §36)

- Vector and tensor indices are denoted by Latin letters i, k, l

- e_{ikl} antisymmetric unit tensor (see §26)

- References to other volumes in the Course of Theoretical Phyncs:

  - *Mechanics* = Vol. 1 (*Mechanics*, third English edition, 1976).
  - *Fields* = Vol. 2 (*The Classical Theory of Fields*, fourth English edition, 1975).
  - *RQT* or *Relativistic Quantum Theory* = Vol. 4 (*Relativistic Quantum Theory*, first English edition, Part 1, 1971; Part 2, 1974); the second English edition appeared in one volume as *Quantum Electrodynamics*, 1982.

All are published by Pergamon Press.

CHAPTER I

THE BASIC CONCEPTS OF QUANTUM MECHANICS

§ 1. The uncertainty principle

When we attempt to apply classical mechanics and electrodynamics to explain atomic phenomena, they lead to results which are in obvious conflict with experiment. This is very clearly seen from the contradiction obtained on applying ordinary electrodynamics to a model of an atom in which the electrons move round the nucleus in classical orbits. During such motion, as in any accelerated motion of charges, the electrons would have to emit electromagnetic waves continually. By this emission, the electrons would lose their energy, and this would eventually cause them to fall into the nucleus. Thus, according to classical electrodynamics, the atom would be unstable, which does not at all agree with reality.

This marked contradiction between theory and experiment indicates that the construction of a theory applicable to atomic phenomena—that is, phenomena occurring in particles of very small mass at very small distances—demands a fundamental modification of the basic physical concepts and laws.

As a starting-point for an investigation of these modifications, it is convenient to take the experimentally observed phenomenon known as *electron diffraction*.[1] It is found that, when a homogeneous beam of electrons passes through a crystal, the emergent beam exhibits a pattern of alternate maxima and minima of intensity, wholly similar to the diffraction pattern observed in the diffraction of electromagnetic waves. Thus, under certain conditions, the behaviour of material particles—in this case, the electrons—displays features belonging to wave processes.

---

[1] The phenomenon of electron diffraction was in fact discovered after quantum mechanics was invented. In our discussion, however, we shall not adhere to the historical sequence of development of the theory, but shall endeavour to construct it in such a way that the connection between the basic principles of quantum mechanics and the experimentally observed phenomena is most clearly shown.

How markedly this phenomenon contradicts the usual ideas of motion is best seen from the following imaginary experiment, an idealization of the experiment of electron diffraction by a crystal. Let us imagine a screen impermeable to electrons, in which two slits are cut. On observing the passage of a beam of electrons[2] through one of the slits, the other being covered, we obtain, on a continuous screen placed behind the slit, some pattern of intensity distribution; in the same way, by uncovering the second slit and covering the first, we obtain another pattern. On observing the passage of the beam through both slits, we should expect, on the basis of ordinary classical ideas, a pattern which is a simple superposition of the other two: each electron, moving in its path, passes through one of the slits and has no effect on the electrons passing through the other slit. The phenomenon of electron diffraction shows, however, that in reality we obtain a diffraction pattern which, owing to interference, does not at all correspond to the sum of the patterns given by each slit separately. It is clear that this result can in no way be reconciled with the idea that electrons move in paths.

Thus the mechanics which governs atomic phenomena—*quantum mechanics* or *wave mechanics*—must be based on ideas of motion which are fundamentally different from those of classical mechanics. In quantum mechanics there is no such concept as the path of a particle. This forms the content of what is called the uncertainty principle, one of the fundamental principles of quantum mechanics, discovered by W. Heisenberg in 1927.[3]

In that it rejects the ordinary ideas of classical mechanics, the uncertainty principle might be said to be negative in content. Of course, this principle in itself does not suffice as a basis on which to construct a new mechanics of particles. Such a theory must naturally be founded on some positive assertions, which we shall discuss below (§2). However, in order to formulate these assertions, we must first ascertain the statement of the problems which confront quantum mechanics. To do so, we first examine the special nature of the interrelation between quantum mechanics and classical mechanics. A more general theory can usually be formulated in a logically complete manner, independently of a less general theory which forms a limiting case of it. Thus, relativistic mechanics can be constructed on the basis of its own fundamental principles, without any reference to Newtonian mechanics. It is in principle impossible, however, to formulate the basic concepts of quantum

---

[2] The beam is supposed so rarefied that the interaction of the particles in it plays no part.

[3] It is of interest to note that the complete mathematical formalism of quantum mechanics was constructed by W. Heisenberg and E. Schrödinger in 1925–6, before the discovery of the uncertainty principle, which revealed the physical content of this formalism.

mechanics without using classical mechanics. The fact that an electron[4] has no definite path means that it has also, in itself, no other dynamical characteristics[5] . Hence it is clear that, for a system composed only of quantum objects, it would be entirely impossible to construct any logically independent mechanics. The possibility of a quantitative description of the motion of an electron requires the presence also of physical objects which obey classical mechanics to a sufficient degree of accuracy. If an electron interacts with such a "classical object", the state of the latter is, generally speaking, altered. The nature and magnitude of this change depend on the state of the electron, and therefore may serve to characterize it quantitatively.

In this connection the "classical object" is usually called *apparatus*, and its interaction with the electron is spoken of as *measurement*. However, it must be emphasized that we are here not discussing a process of measurement in which the physicist-observer takes part. By *measurement*, in quantum mechanics, we understand any process of interaction between classical and quantum objects, occurring apart from and independently of any observer. The importance of the concept of measurement in quantum mechanics was elucidated by N. Bohr.

We have defined "apparatus" as a physical object which is governed, with sufficient accuracy, by classical mechanics. Such, for instance, is a body of large enough mass. However, it must not be supposed that apparatus is necessarily macroscopic. Under certain conditions, the part of apparatus may also be taken by an object which is microscopic, since the idea of "with sufficient accuracy" depends on the actual problem proposed. Thus, the motion of an electron in a Wilson chamber is observed by means of the cloudy track which it leaves, and the thickness of this is large compared with atomic dimensions; when the path is determined with such low accuracy, the electron is an entirely classical object.

Thus quantum mechanics occupies a very unusual place among physical theories: it contains classical mechanics as a limiting case, yet at the same time it requires this limiting case for its own formulation.

We may now formulate the problem of quantum mechanics. A typical problem consists in predicting the result of a subsequent measurement from the known results of previous measurements. Moreover, we shall see later that, in comparison with classical mechanics, quantum mechanics, generally speaking, restricts the range of values which can be taken by various physical

---

[4] In this and the following sections we shall, for brevity, speak of "an electron", meaning in general any object of a quantum nature, i.e. a particle or system of particles obeying quantum mechanics and not classical mechanics.

[5] We refer to quantities which characterize the motion of the electron, and not to those, such as the charge and the mass, which relate to it as a particle; these are parameters.

quantities (for example, energy): that is, the values which can be obtained as a result of measuring the quantity concerned. The methods of quantum mechanics must enable us to determine these admissible values.

The measuring process has in quantum mechanics a very important property: it always affects the electron subjected to it, and it is in principle impossible to make its effect arbitrarily small, for a given accuracy of measurement. The more exact the measurement, the stronger the effect exerted by it, and only in measurements of very low accuracy can the effect on the measured object be small. This property of measurements is logically related to the fact that the dynamical characteristics of the electron appear only as a result of the measurement itself. It is clear that, if the effect of the measuring process on the object of it could be made arbitrarily small, this would mean that the measured quantity has in itself a definite value independent of the measurement.

Among the various kinds of measurement, the measurement of the coordinates of the electron plays a fundamental part. Within the limits of applicability of quantum mechanics, a measurement of the coordinates of an electron can always be performed[6] with any desired accuracy.

Let us suppose that, at definite time intervals Δt, successive measurements of the coordinates of an electron are made. The results will not in general lie on a smooth curve. On the contrary, the more accurately the measurements are made, the more discontinuous and disorderly will be the variation of their results, in accordance with the non-existence of a path of the electron. A fairly smooth path is obtained only if the coordinates of the electron are measured with a low degree of accuracy, as for instance from the condensation of vapour droplets in a Wilson chamber.

If now, leaving the accuracy of the measurements unchanged, we diminish the intervals Δt between measurements, then adjacent measurements, of course, give neighbouring values of the coordinates. However, the results of a series of successive measurements, though they lie in a small region of space, will be distributed in this region in a wholly irregular manner, lying on no smooth curve. In particular, as Δt tends to zero, the results of adjacent measurements by no means tend to lie on one straight line.

This circumstance shows that, in quantum mechanics, there is no such concept as the velocity of a particle in the classical sense of the word, i.e. the limit to which the difference of the coordinates at two instants, divided by the interval Δt between these instants, tends as Δt tends to zero. However,

---

[6]) Once again we emphasize that, in speaking of "performing a measurement", we refer to the interaction of an electron with a classical "apparatus", which in no way presupposes the presence of an external observer.

we shall see later that in quantum mechanics, nevertheless, a reasonable definition of the velocity of a particle at a given instant can be constructed, and this velocity passes into the classical velocity as we pass to classical mechanics. But whereas in classical mechanics a particle has definite coordinates and velocity at any given instant, in quantum mechanics the situation is entirely different. If, as a result of measurement, the electron is found to have definite coordinates, then it has no definite velocity whatever. Conversely, if the electron has a definite velocity, it cannot have a definite position in space. For the simultaneous existence of the coordinates and velocity would mean the existence of a definite path, which the electron has not. Thus, in quantum mechanics, the coordinates and velocity of an electron are quantities which cannot be simultaneously measured exactly, i.e. they cannot simultaneously have definite values. We may say that the coordinates and velocity of the electron are quantities which do not exist simultaneously. In what follows we shall derive the quantitative relation which determines the possibility of an inexact measurement of the coordinates and velocity at the same instant.

A complete description of the state of a physical system in classical mechanics is effected by stating all its coordinates and velocities at a given instant; with these initial data, the equations of motion completely determine the behaviour of the system at all subsequent instants. In quantum mechanics such a description is in principle impossible, since the coordinates and the corresponding velocities cannot exist simultaneously. Thus a description of the state of a quantum system is effected by means of a smaller number of quantities than in classical mechanics, i.e. it is less detailed than a classical description.

A very important consequence follows from this regarding the nature of the predictions made in quantum mechanics. Whereas a classical description suffices to predict the future motion of a mechanical system with complete accuracy, the less detailed description given in quantum mechanics evidently cannot be enough to do this. This means that, even if an electron is in a state described in the most complete manner possible in quantum mechanics, its behaviour at subsequent instants is still in principle uncertain. Hence quantum mechanics cannot make completely definite predictions concerning the future behaviour of the electron. For a given initial state of the electron, a subsequent measurement can give various results. The problem in quantum mechanics consists in determining the probability of obtaining various results on performing this measurement. It is understood, of course, that in some cases the probability of a given result of measurement may be equal to unity, i.e. certainty, so that the result of that measurement is unique.

All measuring processes in quantum mechanics may be divided into two

classes. In one, which contains the majority of measurements, we find those which do not, in any state of the system, lead with certainty to a unique result. The other class contains measurements such that for every possible result of measurement there is a state in which the measurement leads with certainty to that result. These latter measurements, which may be called *predictable*, play an important part in quantum mechanics. The quantitative characteristics of a state which are determined by such measurements are what are called *physical quantities* in quantum mechanics. If in some state a measurement gives with certainty a unique result, we shall say that in this state the corresponding physical quantity has a definite value. In future we shall always understand the expression "physical quantity" in the sense given here.

We shall often find in what follows that by no means every set of physical quantities in quantum mechanics can be measured simultaneously, i.e. can all have definite values at the same time. We have already mentioned one example, namely the velocity and coordinates of an electron. An important part is played in quantum mechanics by sets of physical quantities having the following property: these quantities can be measured simultaneously, but if they simultaneously have definite values, no other physical quantity (not being a function of these) can have a definite value in that state. We shall speak of such sets of physical quantities as *complete sets*.

Any description of the state of an electron arises as a result of some measurement. We shall now formulate the meaning of a *complete description* of a state in quantum mechanics. Completely described states occur as a result of the simultaneous measurement of a complete set of physical quantities. From the results of such a measurement we can, in particular, determine the probability of various results of any subsequent measurement, regardless of the history of the electron prior to the first measurement.

From now on (except in §14) we shall understand by the states of a quantum system just these completely described states.

§ 2. The principle of superposition

The radical change in the physical concepts of motion in quantum mechanics as compared with classical mechanics demands, of course, an equally radical change in the mathematical formalism of the theory. We must therefore consider first of all the way in which states are described in quantum mechanics.

We shall denote by q the set of coordinates of a quantum system, and by $\mathrm{d}q$ the product of the differentials of these coordinates. This $\mathrm{d}q$ is called an

element of volume in the *configuration space* of the system; for one particle, dq coincides with an element of volume dV in ordinary space.

The basis of the mathematical formalism of quantum mechanics lies in the proposition that the state of a system can be described by a definite (in general complex) function $\Psi(q)$ of the coordinates. The square of the modulus of this function determines the probability distribution of the values of the coordinates: $|\Psi|^2 dq$ is the probability that a measurement performed on the system will find the values of the coordinates to be in the element dq of configuration space. The function Ψ is called the *wave function* of the system.[7]

A knowledge of the wave function allows us, in principle, to calculate the probability of the various results of any measurement (not necessarily of the coordinates) also. All these probabilities are determined by expressions bilinear in Ψ and Ψ^*. The most general form of such an expression is

$$\iint \Psi(q)\Psi^*(q')\varphi(q,q')\,dq\,dq', \tag{2.1}$$

where the function $\varphi(q,q)$ depends on the nature and the result of the measurement, and the integration is extended over all configuration space. The probability $\Psi\Psi^*$ of various values of the coordinates is itself an expression of this type.[8]

The state of the system, and with it the wave function, in general varies with time. In this sense the wave function can be regarded as a function of time also. If the wave function is known at some initial instant, then, from the very meaning of the concept of complete description of a state, it is in principle determined at every succeeding instant. The actual dependence of the wave function on time is determined by equations which will be derived later.

The sum of the probabilities of all possible values of the coordinates of the system must, by definition, be equal to unity. It is therefore necessary that the result of integrating $|\Psi|^2$ over all configuration space should be equal to unity:

$$\int |\Psi|^2 dq = 1. \tag{2.2}$$

This equation is what is called the *normalization condition* for wave functions. If the integral of $|\Psi|^2$ converges, then by choosing an appropriate

---

[7] It was first introduced into quantum mechanics by Schrödinger in 1926.

[8] It is obtained from (2.1) when $\varphi(q,q) = \delta(q-q_0)\delta(q-q_0)$, where δ denotes the delta function, defined in §5 below; q_0 denotes the value of the coordinates whose probability is required.

constant coefficient the function Ψ can always be, as we say, *normalized*. However, we shall see later that the integral of $|\Psi|^2$ may diverge, and then Ψ cannot be normalized by the condition (2.2). In such cases $|\Psi|^2$ does not, of course, determine the absolute values of the probability of the coordinates, but the ratio of the values of $|\Psi|^2$ at two different points of configuration space determines the relative probability of the corresponding values of the coordinates.

Since all quantities calculated by means of the wave function, and having a direct physical meaning, are of the form (2.1), in which Ψ appears multiplied by Ψ^*, it is clear that the normalized wave function is determined only to within a constant *phase factor* of the form $e^{i\alpha}$ (where α is any real number). This indeterminacy is in principle irremovable; it is, however, unimportant, since it has no effect upon any physical results.

The positive content of quantum mechanics is founded on a series of propositions concerning the properties of the wave function. These are as follows.

Suppose that, in a state with wave function $\Psi_1(q)$, some measurement leads with certainty to a definite result (result 1), while in a state with $\Psi_2(q)$ it leads to result 2. Then it is assumed that every linear combination of Ψ_1 and Ψ_2, i.e. every function of the form $c_1\Psi_1(q) + c_2\Psi_2(q)$ (where c_1 and c_1 are constants), gives a state in which that measurement leads to either result 1 or result 2. Moreover, we can assert that, if we know the time dependence of the states, which for the one case is given by the function $\Psi_1(q,t)$, and for the other by $\Psi_2(q,t)$, then any linear combination also gives a possible dependence of a state on time. These propositions constitute what is called the *principle of superposition of states*, the chief positive principle of quantum mechanics. In particular, it follows from this principle that all equations satisfied by wave functions must be linear in Ψ.

Let us consider a system composed of two parts, and suppose that the state of this system is given in such a way that each of its parts is completely described.[9]) Then we can say that the probabilities of the coordinates q_1 of the first part are independent of the probabilities of the coordinates q_2 of the second part, and therefore the probability distribution for the whole system should be equal to the product of the probabilities of its parts. This means that the wave function $\Psi_{12}(q_1, q_2)$ of the system can be represented in the

---

[9]) This, of course, means that the state of the whole system is completely described also. However, we emphasize that the converse statement is by no means true: a complete description of the state of the whole system does not in general completely determine the states of its individual parts (see also §14).

form of a product of the wave functions $\Psi_1(q_1)$ and $\Psi_2(q_2)$ of its parts:

$$\Psi_{12}(q_1, q_2) = \Psi_1(q_1)\Psi_2(q_2). \tag{2.3}$$

If the two parts do not interact, then this relation between the wave function of the system and those of its parts will be maintained at future instants also, i.e. we can write

$$\Psi_{12}(q_1, q_2, t) = \Psi_1(q_1, t)\Psi_2(q_2, t). \tag{2.4}$$

§ 3. Operators

Let us consider some physical quantity f which characterizes the state of a quantum system. Strictly, we should speak in the following discussion not of one quantity, but of a complete set of them at the same time. However, the discussion is not essentially changed by this, and for brevity and simplicity we shall work below in terms of only one physical quantity.

The values which a given physical quantity can take are called in quantum mechanics its *eigenvalues*, and the set of these is referred to as the *spectrum of eigenvalues* of the given quantity. In classical mechanics, generally speaking, quantities run through a continuous series of values. In quantum mechanics also there are physical quantities (for instance, the coordinates) whose eigenvalues occupy a continuous range; in such cases we speak of a *continuous spectrum of eigenvalues*. As well as such quantities, however, there exist in quantum mechanics others whose eigenvalues form some discrete set; in such cases we speak of a *discrete spectrum*.

We shall suppose for simplicity that the quantity f considered here has a discrete spectrum; the case of a continuous spectrum will be discussed in §5. The eigenvalues of the quantity f are denoted by f_n, where the suffix n takes the values $0, 1, 2, 3, \ldots$. We also denote the wave function of the system, in the state where the quantity f has the value f_n, by Ψ_n. The wave functions Ψ_n are called the *eigenfunctions* of the given physical quantity f. Each of these functions is supposed normalized, so that

$$\int |\Psi_n|^2 \mathrm{d}q = 1. \tag{3.1}$$

If the system is in some arbitrary state with wave function Ψ, a measurement of the quantity f carried out on it will give as a result one of the eigenvalues f_n. In accordance with the principle of superposition, we can assert that the wave function Ψ must be a linear combination of those eigenfunctions Ψ_n which correspond to the values f_n that can be obtained, with probability

different from zero, when a measurement is made on the system and it is in the state considered. Hence, in the general case of an arbitrary state, the function Ψ can be represented in the form of a series

$$\Psi = \sum a_n \Psi_n, \tag{3.2}$$

where the summation extends over all n, and the a_n are some constant coefficients.

Thus we reach the conclusion that any wave function can be, as we say, expanded in terms of the eigenfunctions of any physical quantity. A set of functions in terms of which such an expansion can be made is called a *complete* (or *closed*) set.

The expansion (3.2) makes it possible to determine the probability of finding (i.e. the probability of getting the corresponding result on measurement), in a system in a state with wave function Ψ, any given value f_n of the quantity f. For, according to what was said in the previous section, these probabilities must be determined by some expressions bilinear in Ψ and Ψ^*, and therefore must be bilinear in a_n and a_n^*. Furthermore, these expressions must, of course, be positive. Finally, the probability of the value f_n must become unity if the system is in a state with wave function $\Psi = \Psi_n$, and must become zero if there is no term containing Ψ_n in the expansion (3.2) of the wave function Ψ. The only essentially positive quantity satisfying these conditions is the square of the modulus of the coefficient a_n. Thus we reach the result that the squared modulus $|a_n|^2$ of each coefficient in the expansion (3.2) determines the probability of the corresponding value f_n of the quantity f in the state with wave function Ψ. The sum of the probabilities of all possible values f_n must be equal to unity; in other words, the relation

$$\sum_n |a_n|^2 = 1 \tag{3.3}$$

must hold.

If the function Ψ were not normalized, then the relation (3.3) would not hold either. The sum $\sum |a_n|^2$ would then be given by some expression bilinear in Ψ and Ψ^*, and becoming unity when Ψ was normalized. Only the integral $\int \Psi \Psi^* \mathrm{d}q$ is such an expression. Thus the equation

$$\sum_n a_n a_n^* = \int \Psi \Psi^* \mathrm{d}q \tag{3.4}$$

must hold.

On the other hand, multiplying by Ψ the expansion of the function Ψ^* (the complex conjugate of Ψ), and integrating, we obtain

$$\int \Psi\Psi^* \mathrm{d}q = \sum_n a_n^* \int \Psi_n^* \Psi \mathrm{d}q.$$

Comparing this with (3.4), we have

$$\sum_n a_n a_n^* = \sum_n a_n^* \int \Psi_n^* \Psi \mathrm{d}q,$$

from which we derive the following formula determining the coefficients a_n in the expansion of the function Ψ in terms of the eigenfunctions Ψ_n:

$$a_n = \int \Psi\Psi_n^* \mathrm{d}q. \tag{3.5}$$

If we substitute here from (3.2), we obtain

$$a_n = \sum_m a_m \int \Psi_m \Psi_n^* \mathrm{d}q$$

from which it is evident that the eigenfunctions must satisfy the conditions

$$\int \Psi_m \Psi_n^* \mathrm{d}q = \delta_{nm} \tag{3.6}$$

where $\delta_{nm} = 1$ for $n = m$ and $\delta_{nm} = 0$ for $n \neq m$. The fact that the integrals of the products $\Psi_m \Psi_n^*$ with $n \neq m$ vanish is called the *orthogonality* of the functions Ψ_n. Thus the set of eigenfunctions Ψ_n forms a complete set of normalized and orthogonal (or, for brevity, *orthonormal*) functions.

We shall now introduce the concept of the mean value $\overline{f}$ of the quantity f in the given state. In accordance with the usual definition of mean values, we define $\overline{f}$ as the sum of all the eigenvalues f_n of the given quantity, each multiplied by the corresponding probability $|a_n|^2$. Thus

$$\overline{f} = \sum_n f_n |a_n|^2. \tag{3.7}$$

We shall write $\overline{f}$ in the form of an expression which does not contain the coefficients a_n in the expansion of the function Ψ, but this function itself. Since the products $a_n a_n*$ appear in (3.7), it is clear that the required expression must be bilinear in Ψ and Ψ^*. We introduce a mathematical operator,

which we denote[10]) by $\widehat{f}$ and define as follows. Let $(\widehat{f}\Psi)$ denote the result of the operator $\widehat{f}$ acting on the function Ψ. We define $\widehat{f}$ in such a way that the integral of the product of $(\widehat{f}\Psi)$ and the complex conjugate function Ψ^* is equal to the mean value $\overline{f}$:

$$\overline{f} = \int \Psi^*(\widehat{f}\Psi)\mathrm{d}q. \tag{3.8}$$

It is easily seen that, in the general case, the operator $\widehat{f}$ is a linear[11]) integral operator. For, using the expression (3.5) for a_n, we can rewrite the definition (3.7) of the mean value in the form

$$\overline{f} = \sum_n f_n a_n a_n^* = \int \Psi^*(\sum_n a_n f_n \Psi_n)\mathrm{d}q.$$

Comparing this with (3.8), we see that the result of the operator $\widehat{f}$ acting on the function Ψ has the form

$$(\widehat{f}\Psi) = \sum_n a_n f_n \Psi_n. \tag{3.9}$$

If we substitute here the expression (3.5) for a_n, we find that $\widehat{f}$ is an integral operator of the form

$$(\widehat{f}\Psi) = \int K(q, q')\Psi(q')\mathrm{d}q', \tag{3.10}$$

where the function $K(q, q')$ (called the *kernel* of the operator) is

$$K(q, q') = \sum_n f_n \Psi_n^*(q')\psi_n(q). \tag{3.11}$$

Thus, for every physical quantity in quantum mechanics, there is a definite corresponding linear operator.

It is seen from (3.9) that, if the function Ψ is one of the eigenfunctions Ψ_n(so that all the an except one are zero), then, when the operator $\widehat{f}$ acts on it, this function is simply multiplied by the corresponding eigenvalue f_n:

$$\widehat{f}\Psi_n = f_n\Psi_n. \tag{3.12}$$

---

[10]) By convention, we shall always denote operators by letters with circumflexes.

[11]) An operator is said to be linear if it has the properties

$$\widehat{f}(\Psi_1 + \Psi_2) = \widehat{f}(\Psi_1) + \widehat{f}(\Psi_2) \text{ and } \widehat{f}(a\Psi) = a\widehat{f}\Psi$$

where Ψ_1 and Ψ_2 are arbitrary functions and a is an arbitrary constant.

(In what follows we shall always omit the parentheses in the expression $(\widehat{f}\Psi)$, where this cannot cause any misunderstanding; the operator is taken to act on the expression which follows it.) Thus we can say that the eigenfunctions of the given physical quantity f are the solutions of the equation

$$\widehat{f}\Psi = f\Psi,$$

where f is a constant, and the eigenvalues are the values of this constant for which the above equation has solutions satisfying the required conditions. As we shall see below, the form of the operators for various physical quantities can be determined from direct physical considerations, and then the above property of the operators enables us to find the eigenfunctions and eigenvalues by solving the equations $\widehat{f}\Psi = f\Psi$.

Both the eigenvalues of a real physical quantity and its mean value in every state are real. This imposes a restriction on the corresponding operators. Equating the expression (3.8) to its complex conjugate, we obtain the relation

$$\int \Psi^*(\widehat{f}\Psi)\mathrm{d}q = \int \psi(\widehat{f}^*\Psi^*)\mathrm{d}q, \tag{3.13}$$

where $\widehat{f}^*$ denotes the operator which is the complex conjugate of $\widehat{f}$.[12] This relation does not hold in general for an arbitrary linear operator, so that it is a restriction on the form of the operator $\widehat{f}$. For an arbitrary operator $\widehat{f}$ we can find what is called the *transposed operator* $\widetilde{f}$, defined in such a way that

$$\int \Phi(\widehat{f}^*\Psi)\mathrm{d}q = \int \Psi(\widetilde{\widehat{f}}\Phi)\mathrm{d}q \tag{3.14}$$

where Ψ and Φ are two different functions. If we take, as the function Φ, the function Ψ^* which is the complex conjugate of Ψ, then a comparison with (3.13) shows that we must have

$$\widetilde{\widehat{f}} = \widehat{f}^* \tag{3.15}$$

Operators satisfying this condition are said to be *Hermitian*.[13] Thus the operators corresponding, in the mathematical formalism of quantum mechanics, to real physical quantities must be Hermitian.

We can formally consider complex physical quantities also, i.e. those whose eigenvalues are complex. Let f be such a quantity. Then we can

---

[12] By definition, if for the operator $\widehat{f}$ we have $\widehat{f}\Psi = \varphi$, then the complex conjugate operator $\widehat{f}^*$ is that for which we have $\widehat{f}^*\Psi^* = \varphi^*$.

[13] For a linear integral operator of the form (3.10), the Hermitian condition means that the kernel of the operator must be such that $K(q, q) = K^*(q, q)$.

introduce its complex conjugate quantity f^*, whose eigenvalues are the complex conjugates of those of f. We denote by $\widehat{f}^\dagger$ the operator corresponding to the quantity f^*. It is called the *Hermitian conjugate* of the operator $\widehat{f}$ and, in general, will be different from the definition of the operator $\widehat{f}^*$: the mean value of the quantity f^* in a state Ψ is

$$\overline{f^*} = \int \Psi^* \widehat{f}^\dagger \Psi \mathrm{d}q.$$

We also have

$$(\overline{f})^* = \left[\int \Psi^* \widehat{f} \Psi \mathrm{d}q\right]^*$$

$$= \int \Psi \widehat{f}^* \Psi^* \mathrm{d}q$$

$$= \int \Psi^* \widetilde{\widehat{f}}^* \Psi \mathrm{d}q.$$

Equating these two expressions gives

$$\widehat{f}^\dagger = \widetilde{\widehat{f}}^*, \tag{3.16}$$

from which it is clear that $\widehat{f}^\dagger$ is in general not the same as $\widehat{f}^*$.

The condition (3.15) can now be written

$$\widehat{f} = \widehat{f}^\dagger, \tag{3.17}$$

i.e. the operator of a real physical quantity is the same as its Hermitian conjugate (Hermitian operators are also called *self-conjugate*).

We shall show how the orthogonality of the eigenfunctions of an Hermitian operator corresponding to different eigenvalues can be directly proved. Let f_n and f_m be two different eigenvalues of the real quantity f, and Ψ_n, Ψ_m the corresponding eigenfunctions:

$$\widehat{f}\Psi_n = f_n\Psi_n, \widehat{f}\Psi_m = f_m\Psi_m.$$

Multiplying both sides of the first of these equations by Ψ_m^*, and both sides of the complex conjugate of the second by Ψ_n, and subtracting corresponding terms, we find

$$\Psi_m^* \widehat{f}\Psi_n - \Psi_n \widehat{f}^* \Psi_m^* = (f_n - f_m)\Psi_n\Psi_m^*.$$

We integrate both sides of this equation over q. Since $\widehat{f}^* = \widetilde{\widehat{f}}$, by (3.14) the integral on the left-hand side of the equation is zero, so that we have

$$(f_n - f_m) \int \Psi_n\Psi_m^* \mathrm{d}q = 0$$

whence, since $f_n \neq f_m$, we obtain the required orthogonality property of the functions Ψ_n and Ψ_m.

We have spoken here of only one physical quantity f, whereas, as we said at the beginning of this section, we should have spoken of a complete set of simultaneously measurable physical quantities. We should then have found that to each of these quantities $f, g, \dots$ there corresponds its operator $\widehat{f}, \widehat{g}, \dots$. The eigenfunctions Ψ_n then correspond to states in which all the quantities concerned have definite values, i.e. they correspond to definite sets of eigenvalues $f_n, g_n, \dots$, and are simultaneous solutions of the system of equations

$$\widehat{f}\Psi = f\Psi, \quad \widehat{g}\Psi = g\Psi, \dots.$$

§ 4. Addition and multiplication of operators

If $\widehat{f}$ and $\widehat{g}$ are the operators corresponding to two physical quantities f and g, the sum $f + g$ has a corresponding operator $\widehat{f} + \widehat{g}$. However, the significance of adding different physical quantities in quantum mechanics depends considerably on whether the quantities are or are not simultaneously measurable. If f and g are simultaneously measurable, the operators $\widehat{f}$ and $\widehat{g}$ have common eigenfunctions, which are also eigenfunctions of the operator $\widehat{f} + \widehat{g}$, and the eigenvalues of the latter operator are equal to the sums $f_n + g_n$. But if f and g cannot simultaneously take definite values, their sum $f + g$ has a more restricted significance. We can assert only that the mean value of this quantity in any state is equal to the sum of the mean values of the separate quantities:

$$\overline{f + g} = \overline{f} + \overline{g}. \tag{4.1}$$

The eigenvalues and eigenfunctions of the operator $\widehat{f} + \widehat{g}$ will not, in general, now bear any relation to those of the quantities f and g. It is evident that, if the operators $\widehat{f}$ and $\widehat{g}$ are Hermitian, the operator $\widehat{f} + \widehat{g}$ will be so too, so that its eigenvalues are real and are equal to those of the new quantity $f + g$ thus defined.

The following theorem should be noted. Let f_0 and g_0 be the smallest eigenvalues of the quantities f and g, and $(f + g)_0$ that of the quantity $f + g$. Then

$$(f + g)_0 \geqslant f_0 + g_0 \tag{4.2}$$

The equality holds if f and g can be measured simultaneously. The proof follows from the obvious fact that the mean value of a quantity is always greater than or equal to its least eigenvalue. In a state in which the quantity

$f + g$ has the value $(f + g)_0$ we have $\overline{f + g} = (f + g)_0$, and since, on the other hand, $\overline{f + g} = \overline{f} + \overline{g} \geqslant f_0 + g_0$, we arrive at the inequality (4.2).

Next, let f and g once more be quantities that can be measured simultaneously. Besides their sum, we can also introduce the concept of their *product* as being a quantity whose eigenvalues are equal to the products of those of the quantities f and g. It is easy to see that, to this quantity, there corresponds an operator whose effect consists of the successive action on the function of first one and then the other operator. Such an operator is represented mathematically by the product of the operators $\widehat{f}$ and $\widehat{g}$. For, if Ψ_n are the eigenfunctions common to the operators $\widehat{f}$ and $\widehat{g}$, we have

$$\widehat{f}\widehat{g}\Psi_n = \widehat{f}(\widehat{g}\Psi_n) = \widehat{f}g_n\Psi_n = g_n\widehat{f}\Psi_n = g_nf_n\Psi_n$$

(the symbol $\widehat{f}\widehat{g}$ denotes an operator whose effect on a function Ψ consists of the successive action first of the operator $\widehat{g}$ on the function Ψ and then of the operator $\widehat{f}$ on the function $\widehat{g}\Psi$). We could equally well take the operator $\widehat{g}\widehat{f}$ instead of $\widehat{f}\widehat{g}$, the former differing from the latter in the order of its factors. It is obvious that the result of the action of either of these operators on the functions Ψ_n will be the same. Since, however, every wave function Ψ can be represented as a linear combination of the functions Ψ_n, it follows that the result of the action of the operators $\widehat{f}\widehat{g}$ and $\widehat{g}\widehat{f}$ on an arbitrary function will also be the same. This fact can be written in the form of the symbolic equation $\widehat{f}\widehat{g} = \widehat{g}\widehat{f}$ or

$$\widehat{f}\widehat{g} - \widehat{g}\widehat{f} = 0 \tag{4.3}$$

Two such operators $\widehat{f}$ and $\widehat{g}$ are said to be *commutative*, or to *commute* with each other. Thus we arrive at the important result: if two quantities f and g can simultaneously take definite values, then their operators commute with each other.

The converse theorem can also be proved (§11): if the operators $\widehat{f}$ and $\widehat{g}$ commute, then all their eigenfunctions can be taken common to both; physically, this means that the corresponding physical quantities can be measured simultaneously. Thus the commutability of the operators is a necessary and sufficient condition for the physical quantities to be simultaneously measurable.

A particular case of the product of operators is an operator raised to some power. From the above discussion we can deduce that the eigenvalues of an operator $\widehat{f}^p$ (where p is an integer) are equal to the pth powers of the eigenvalues of the operator $\widehat{f}$. Any function $\varphi(\widehat{f})$ of an operator can be defined as an operator whose eigenvalues are equal to the same function $\varphi(f)$ of the eigenvalues of the operator $\widehat{f}$. If the function $\varphi(f)$ can be expanded

as a Taylor series, this expresses the effect of the operator $\varphi(\widehat{f})$ in terms of those of various powers $\widehat{f}^p$.

In particular, the operator $\widehat{f}^{-1}$ is called the inverse of the operator $\widehat{f}$. It is evident that the successive action of the operators $\widehat{f}$ and $\widehat{f}^{-1}$ on any function leaves the latter unchanged, i.e. $\widehat{f}\widehat{f}^{-1} = \widehat{f}^{-1}\widehat{f} = 1$.

If the quantities f and g cannot be measured simultaneously, the concept of their product does not have the same direct meaning. This appears in the fact that the operator $\widehat{fg}$ is not Hermitian in this case, and hence cannot correspond to any real physical quantity. For, by the definition of the transpose of an operator we can write

$$\int \Psi \widehat{fg} \Phi \mathrm{d}q = \int \Psi \widehat{f}(\widehat{g}\Phi)\mathrm{d}q = \int (\widehat{g}\Phi)(\widetilde{\widehat{f}}\Psi)\mathrm{d}q.$$

Here the operator $\widetilde{\widehat{f}}$ acts only on the function Ψ, and the operator $\widehat{g}$ on Φ, so that the integrand is a simple product of two functions $\widehat{g}\Phi$ and $\widetilde{\widehat{f}}\Psi$. Again using the definition of the transpose of an operator, we can write

$$\int \Psi \widehat{fg} \Phi \mathrm{d}q = \int (\widetilde{\widehat{f}}\Psi)(\widehat{g}\Phi)\mathrm{d}q = \int \Phi \widetilde{\widehat{g}}\widetilde{\widehat{f}}\Phi \mathrm{d}q.$$

Thus we obtain an integral in which the functions Ψ and Φ have changed places as compared with the original one. In other words, the operator $\widetilde{\widehat{g}}\widetilde{\widehat{f}}$ is the transpose of $\widehat{fg}$, and we can write

$$\widetilde{\widehat{fg}} = \widetilde{\widehat{g}}\widetilde{\widehat{f}} \tag{4.4}$$

i.e. the transpose of the product $\widehat{fg}$ is the product of the transposes of the factors written in the opposite order. Taking the complex conjugate of both sides of equation (4.4), we have

$$(\widehat{fg})^\dagger = \widehat{g}^\dagger \widehat{f}^\dagger \tag{4.5}$$

If each of the operators $\widehat{f}$ and $\widehat{g}$ is Hermitian, then $(\widehat{fg})^\dagger = \widehat{g}\widehat{f}$. It follows from this that the operator $\widehat{fg}$ is Hermitian if and only if the factors $\widehat{f}$ and $\widehat{g}$ commute.

We note that, from the products $\widehat{fg}$ and $\widehat{g}\widehat{f}$ of two non-commuting Hermitian operators, we can form an Hermitian operator, the *symmetrized product*

$$\frac{1}{2}(\widehat{fg} + \widehat{g}\widehat{f}) \tag{4.6}$$

It is easy to see that the difference $\widehat{f}\widehat{g} - \widehat{g}\widehat{f}$ is an *anti-Hermitian* operator (i.e. one for which the transpose is equal to the complex conjugate taken with the opposite sign). It can be made Hermitian by multiplying by i; thus

$$\mathrm{i}(\widehat{f}\widehat{g} - \widehat{g}\widehat{f}) \tag{4.7}$$

is again an Hermitian operator.

In what follows we shall sometimes use for brevity the notation

$$\{\widehat{f}, \widehat{g}\} = \widehat{f}\widehat{g} - \widehat{g}\widehat{f} \tag{4.8}$$

called the commutator of these operators. It is easily seen that

$$\{\widehat{f}\widehat{g}, \widehat{h}\} = \{\widehat{f}, \widehat{h}\}\widehat{g} + \widehat{f}\widehat{g}, \widehat{h} \tag{4.9}$$

We notice that, if $\{\widehat{f}, \widehat{h}\} = 0$ and $\{\widehat{g}, \widehat{h}\} = 0$, it does not in general follow that $\widehat{f}$ and $\widehat{g}$ commute.

§ 5. The continuous spectrum

All the relations given in §3 and 4, describing the properties of the eigenfunctions of a discrete spectrum, can be generalized without difficulty to the case of a continuous spectrum of eigenvalues.

Let f be a physical quantity having a continuous spectrum. We shall denote its eigenvalues by the same letter f simply, and the corresponding eigenfunctions by Ψ_f. Just as an arbitrary wave function Ψ can be expanded in a series (3.2) of eigenfunctions of a quantity having a discrete spectrum, it can also be expanded (this time as an integral) in terms of the complete set of eigenfunctions of a quantity with a continuous spectrum. This expansion has the form

$$\Psi(q) = \int a_f \Psi_f(q)\mathrm{d}f \tag{5.1}$$

where the integration is extended over the whole range of values that can be taken by the quantity f.

The subject of the normalization of the eigenfunctions of a continuous spectrum is more complex than in the case of a discrete spectrum. The requirement that the integral of the squared modulus of the function should be equal to unity cannot here be satisfied, as we shall see below. Instead, we try to normalize the functions Ψ_f in such a way that $|a_f|^2\mathrm{d}f$ is the probability that the physical quantity concerned, in the state described by the

wave function Ψ, has a value between f and $f + \mathrm{d}f$. Since the sum of the probabilities of all possible values of f must be equal to unity, we have

$$\int |a_f|^2 \mathrm{d}f = 1 \tag{5.2}$$

(similarly to the relation (3.3) for a discrete spectrum).

Proceeding in exactly the same way as in the derivation of formula (3.5), and using the same arguments, we can write, firstly,

$$\int \Psi\Psi^* \mathrm{d}q = \int |a_f|^2 \mathrm{d}f$$

and, secondly,

$$\int \Psi\Psi^* \mathrm{d}q = \iint a_f^* \Psi_f^* \Psi \mathrm{d}f \mathrm{d}q.$$

By comparing these two expressions we find the formula which determines the expansion coefficients,

$$a_f = \int \Psi(q)\Psi_f^*(q)\mathrm{d}q, \tag{5.3}$$

in exact analogy to (3.5).

To derive the normalization condition, we now substitute (5.1) in (5.3), and obtain

$$a_f = \int a_{f'} \left(\int \Psi_{f'}\Psi_f^* \mathrm{d}q \right) \mathrm{d}f'$$

This relation must hold for arbitrary a_f, and therefore must be satisfied identically. For this to be so, it is necessary that, first of all, the coefficient of a_f, in the integrand (i.e. the integral $\int \Psi_{f'}\Psi_f^* \mathrm{d}q$) should be zero for all $f \neq f$. For $f = f$, this coefficient must become infinite (otherwise the integral over f would vanish). Thus the integral $\int \Psi_{f'}\Psi_f^* \mathrm{d}q$ is a function of the difference $f' - f$, which becomes zero for values of the argument different from zero and is infinite when the argument is zero. We denote this function by $\delta(f' - f)$:

$$\int \Psi_{f'}\Psi_f^* \mathrm{d}q = \delta(f' - f) \tag{5.4}$$

The manner in which the function $\delta(f' - f)$ becomes infinite for $f' - f = 0$ is determined by the fact that we must have

$$\int \delta(f' - f)a_{f'}\mathrm{d}f' = a_f$$

It is clear that, for this to be so, we must have

$$\int \delta(f' - f)\mathrm{d}f' = 1$$

The function thus defined is called a *delta function*, and was first used in theoretical physics by P. A. M. Dirac. We shall write out once more the formulae which define it. They are

$$\delta(x) = 0 \text{ for } x \neq 0, \qquad \delta(0) = \infty, \tag{5.5}$$

while

$$\int_{-\infty}^{+\infty} \delta(x)\mathrm{d}x = 1. \tag{5.6}$$

We can take as limits of integration any numbers such that $x = 0$ lies between them. If $f(x)$ is some function continuous at $x = 0$, then

$$\int_{-\infty}^{+\infty} \delta(x)f(x)\mathrm{d}x = f(0) \tag{5.7}$$

This formula can be written in the more general form

$$\int \delta(x - a)f(x)\mathrm{d}x = f(a) \tag{5.8}$$

where the range of integration includes the point $x = a$, and $f(x)$ is continuous at $x = a$. It is also evident that

$$\delta(-x) = \delta(x) \tag{5.9}$$

i.e. the delta function is even. Finally, writing

$$\int_{-\infty}^{+\infty} \delta(\alpha x)\mathrm{d}x = \int_{-\infty}^{+\infty} \delta(y)\frac{\mathrm{d}y}{|\alpha|} = \frac{1}{|\alpha|}$$

we can deduce that

$$\delta(\alpha x) = \frac{1}{|\alpha|}\delta(x) \tag{5.10}$$

where α is any constant.

The formula (5.4) gives the normalization rule for the eigenfunctions of a continuous spectrum; it replaces the condition (3.6) for a discrete spectrum. We see that the functions Ψ_f and $\Psi_{f'}$ with $f \neq f'$ are, as before, orthogonal. However, the integrals of the squared moduli $|\Psi_f|^2$ of the functions diverge for a continuous spectrum.

The functions $\Psi_f(q)$ satisfy still another relation similar to (5.4). To derive this, we substitute (5.3) in (5.1), which gives

$$\Psi(q) = \int \Psi(q') \left(\int \Psi_f^*(q')\Psi_f(q)\mathrm{d}f \right) \mathrm{d}q'$$

whence we can at once deduce that we must have

$$\int \Psi_f^*(q')\Psi_f(q)\mathrm{d}f = \delta(q' - q) \tag{5.11}$$

There is, of course, an analogous relation for a discrete spectrum:

$$\sum_n \Psi_n^*(q')\Psi_n(q) = \delta(q' - q) \tag{5.12}$$

Comparing the pair of formulae (5.1), (5.4) with the pair (5.3), (5.11), we see that, on the one hand, the function $\Psi(q)$ can be expanded in terms of the functions$\Psi_f(q)$ with expansion coefficients a_f and, on the other hand, formula (5.3) represents an entirely analogous expansion of the function $a_f \equiv a(f)$ in terms of the functions $\Psi_f^*(q)$, while the $\Psi(q)$ play the part of expansion coefficients. The function $a(f)$, like $\Psi(q)$, completely determines the state of the system; it is sometimes called a wave function *in the f representation* (while the function $\Psi(q)$ is called a wave function in the *q representation*). Just as $|\Psi(q)|^2$ determines the probability for the system to have coordinates lying in a given interval $\mathrm{d}q$, so $|a(f)|^2$ determines the probability for the values of the quantity f to lie in a given interval $\mathrm{d}f$. On the one hand, the functions $\Psi_f(q)$ are the eigenfunctions of the quantity f in the q representation; on the other hand, their complex conjugates are the eigenfunctions of the coordinate q in the f representation.

Let $\varphi(f)$ be some function of the quantity f, such that φ and f are related in a one-to-one manner. Each of the functions $\Psi_f(q)$ can then be regarded as an eigenfunction of the quantity φ. Here, however, the normalization of these functions must be changed: the eigenfunctions $\Psi_\varphi(q)$ of the quantity must be normalized by the condition

$$\int \Psi_{\varphi(f')}\Psi_{\varphi(f)}^* \mathrm{d}q = \delta\left[\varphi(f') - \varphi(f)\right],$$

whereas the functions Ψ_f, are normalized by the condition (5.4). The argument of the delta function becomes zero only for $f' = f$. As f' approaches f, we have

$$\varphi(f') - \varphi(f) = \frac{\mathrm{d}\varphi(f)}{\mathrm{d}f}(f' - f).$$

By (5.10) we can therefore write[14]

$$\delta\left[\varphi(f') - \varphi(f)\right] = \frac{1}{|\mathrm{d}\varphi(f)/\mathrm{d}f|}\delta(f' - f) \tag{5.13}$$

Comparing this with (5.4), we see that the functions Ψ_φ and Ψ_f are related by

$$\Psi_{\varphi(f)} = \frac{1}{\sqrt{|\mathrm{d}\varphi(f)/\mathrm{d}f|}}\Psi_f \tag{5.14}$$

There are also physical quantities which in one range of values have a discrete spectrum, and in another a continuous spectrum. For the eigenfunctions of such a quantity all the relations derived in this and the previous sections are, of course, true. It need only be noted that the complete set of functions is formed by combining the eigenfunctions of both spectra. Hence the expansion of an arbitrary wave function in terms of the eigenfunctions of such a quantity has the form

$$\Psi(q) = \sum_n a_n \Psi_n(q) + \int a_f \Psi_f(q)\mathrm{d}f \tag{5.15}$$

where the sum is taken over the discrete spectrum and the integral over the whole continuous spectrum.

The coordinate q itself is an example of a quantity having a continuous spectrum. It is easy to see that the operator corresponding to it is simply multiplication by q. For, since the probability of the various values of the coordinate is determined by the square $|\Psi(q)|^2$, the mean value of the coordinate is

$$\bar{q} = \int q|\Psi|^2\mathrm{d}q = \int \Psi^* q\Psi\mathrm{d}q.$$

Comparison of this with the definition (3.8) of an operator shows that[15]

$$\widehat{q} = q. \tag{5.16}$$

The eigenfunctions of this operator must be determined, according to the usual rule, by the equation $q\Psi_{q_0} = q_0\Psi_{q_0}$, where q_0 temporarily denotes the

---

[14] In general, if $\varphi(x)$ is some one-valued function (the inverse function need not be one-valued), we have

$$\delta\left[\varphi(x)\right] = \sum_i \frac{1}{|\varphi'(\alpha_i)|}\delta(x - \alpha_i) \tag{5.13a}$$

where α_i are the roots of the equation $\varphi(x) = 0$.

[15] In future we shall always, for simplicity, write operators which amount to multiplication by some quantity in the form of that quantity itself.

actual values of the coordinate as distinct from the variable q. Since this equation can be satisfied either by $\Psi_{q_0} = 0$ or by $q = q_0$, it is clear that the eigenfunctions which satisfy the normalization condition are[16])

$$\Psi_{q_0} = \delta(q - q_0) \tag{5.17}$$

§ 6. The passage to the limiting case of classical mechanics

Quantum mechanics contains classical mechanics in the form of a certain limiting case. The question arises as to how this passage to the limit is made.

In quantum mechanics an electron is described by a wave function which determines the various values of its coordinates; of this function we so far know only that it is the solution of a certain linear partial differential equation. In classical mechanics, on the other hand, an electron is regarded as a material particle, moving in a path which is completely determined by the equations of motion. There is an interrelation, somewhat similar to that between quantum and classical mechanics, in electrodynamics between wave optics and geometrical optics. In wave optics, the electromagnetic waves are described by the electric and magnetic field vectors, which satisfy a definite system of linear differential equations, namely Maxwell's equations. In geometrical optics, however, the propagation of light along definite paths, or rays, is considered. Such an analogy enables us to see that the passage from quantum mechanics to the limit of classical mechanics occurs similarly to the passage from wave optics to geometrical optics.

Let us recall how this latter transition is made mathematically (see Fields, §53). Let u be any of the field components in the electromagnetic wave. It can be written in the form $u = ae^{i\varphi}$ (with a and φ real), where a is called the amplitude and φ the phase of the wave (called in geometrical optics the eikonal). The limiting case of geometrical optics corresponds to small wavelengths; this is expressed mathematically by saying that φ varies by a large amount over short distances; this means, in particular, that it can be

---

[16]) The expansion coefficients for an arbitrary function Ψ in terms of these eigenfunctions are

$$a_{q_0} = \int \Psi(q)\delta(q - q_0)\mathrm{d}q = \Psi(q_0).$$

The probability that the value of the coordinate lies in a given interval $\mathrm{d}q_0$ is

$$|a_{q_0}|^2\mathrm{d}q_0 = |\Psi(q_0)|^2\mathrm{d}q_0,$$

as it should be.

supposed large in absolute value.

Similarly, we start from the hypothesis that, to the limiting case of classical mechanics, there correspond in quantum mechanics wave functions of the form $\Psi = a\mathrm{e}^{\mathrm{i}\varphi}$, where a is a slowly varying function and φ takes large values. As is well known, the path of a particle can be determined in mechanics by means of the variational principle, according to which what is called the action S of a mechanical system must take its least possible value (the principle of least action). In geometrical optics the path of the rays is determined by what is called Fermat's principle, according to which the optical path length of the ray, i.e. the difference between its phases at the beginning and end of the path, must take its least (or greatest) possible value.

On the basis of this analogy, we can assert that the phase φ of the wave function, in the limiting (classical) case, must be proportional to the mechanical action S of the physical system considered, i.e. we must have $S = \mathrm{const}\,\varphi$. The constant of proportionality is called Planck's constant[17] and is denoted by $\hbar$. It has the dimensions of action (since φ is dimensionless) and has the value

$$\hbar = 1.054 \times 10^{-27} \mathrm{erg} \cdot \mathrm{s}$$

Thus, the wave function of an "almost classical" (or, as we say, *quasi-classical*) physical system has the form

$$\Psi = a\mathrm{e}^{\mathrm{i}S/\hbar} \tag{6.1}$$

Planck's constant $\hbar$ plays a fundamental part in all quantum phenomena. Its relative value (compared with other quantities of the same dimensions) determines the "extent of quantization" of a given physical system. The transition from quantum mechanics to classical mechanics, corresponding to large phase, can be formally described as a passage to the limit $\hbar \to 0$ (just as the transition from wave optics to geometrical optics corresponds to a passage to the limit of zero wavelength, $\lambda \to 0$).

We have ascertained the limiting form of the wave function, but the question still remains how it is related to classical motion in a path. In general, the motion described by the wave function does not tend to motion in a definite path. Its connection with classical motion is that, if at some initial instant the wave function, and with it the probability distribution of the coordinates, is given, then at subsequent instants this distribution will change according to the laws of classical mechanics (for a more detailed discussion of this, see the end of §17).

---

[17] It was introduced into physics by M. Planck in 1900. The constant $\hbar$ which we use everywhere in this book, is, strictly speaking, Planck's constant divided by 2π; this is Dirac's notation.

In order to obtain motion in a definite path, we must start from a wave function of a particular form, which is perceptibly different from zero only in a very small region of space (what is called a wave packet); the dimensions of this region must tend to zero with $\hbar$. Then we can say that, in the quasi-classical case, the wave packet will move in space along a classical path of a particle.

Finally, quantum-mechanical operators must reduce, in the limit, simply to multiplication by the corresponding physical quantity.

§ 7. The wave function and measurements

Let us again return to the process of measurement, whose properties have been qualitatively discussed in §1; we shall show how these properties are related to the mathematical formalism of quantum mechanics.

We consider a system consisting of two parts: a classical apparatus and an electron (regarded as a quantum object). The process of measurement consists in these two parts' coming into interaction with each other, as a result of which the apparatus passes from its initial state into some other; from this change of state we draw conclusions concerning the state of the electron. The states of the apparatus are distinguished by the values of some physical quantity (or quantities) characterizing it—the "readings of the apparatus". We conventionally denote this quantity by g, and its eigenvalues by g_n; these take in general, in accordance with the classical nature of the apparatus, a continuous range of values, but we shall—merely in order to simplify the subsequent formulae—suppose the spectrum discrete. The states of the apparatus are described by means of quasi-classical wave functions, which we shall denote by $\Psi_n(\xi)$, where the suffix n corresponds to the "reading" g_n of the apparatus, and ξ denotes the set of its coordinates. The classical nature of the apparatus appears in the fact that, at any given instant, we can say with certainty that it is in one of the known states Ψ_n with some definite value of the quantity g; for a quantum system such an assertion would, of course, be unjustified.

Let $\Phi_0(\xi)$ be the wave function of the initial state of the apparatus (before the measurement), and $\Psi(q)$ some arbitrary normalized initial wave function of the electron (q denoting its coordinates). These functions describe the state of the apparatus and of the electron independently, and therefore the initial wave function of the whole system is the product

$$\Psi(q)\Phi_0(\xi) \tag{7.1}$$

Next, the apparatus and the electron interact with each other. Applying the

equations of quantum mechanics, we can in principle follow the change of the wave function of the system with time. After the measuring process it may not, of course, be a product of functions of ξ and q. Expanding the wave function in terms of the eigenfunctions Φ_n of the apparatus (which form a complete set of functions), we obtain a sum of the form

$$\sum_n A_n(q)\Psi_n(\xi) \tag{7.2}$$

where the $A_n(q)$ are some functions of q.

The classical nature of the apparatus, and the double role of classical mechanics as both the limiting case and the foundation of quantum mechanics, now make their appearance. As has been said above, the classical nature of the apparatus means that, at any instant, the quantity g (the "reading of the apparatus") has some definite value. This enables us to say that the state of the system apparatus + electron after the measurement will in actual fact be described, not by the entire sum (7.2), but by only the one term which corresponds to the "reading" gn of the apparatus,

$$A_n(q)\Psi_n(\xi) \tag{7.3}$$

It follows from this that $A_n(q)$ is proportional to the wave function of the electron after the measurement. It is not the wave function itself, as is seen from the fact that the function $A_n(q)$ is not normalized. It contains both information concerning the properties of the resulting state of the electron and the probability (determined by the initial state of the system) of the occurrence of the nth "reading" of the apparatus.

Since the equations of quantum mechanics are linear, the relation between $A_n(q)$ and the initial wave function of the electron $\Psi(q)$ is in general given by some linear integral operator:

$$A_n(q) = \int K_n(q, q')\Psi(q')\mathrm{d}q' \tag{7.4}$$

with a kernel $K_n(q, q')$ which characterizes the measurement process concerned.

We shall suppose that the measurement concerned is such that it gives a complete description of the state of the electron. In other words (see §1), in the resulting state the probabilities of all the quantities must be independent of the previous state of the electron (before the measurement). Mathematically, this means that the form of the functions $A_n(q)$ must be determined by the measuring process itself, and does not depend on the initial wave function $\Psi(q)$ of the electron. Thus the A_n must have the form

$$A_n(q) = a_n\varphi_n(q) \tag{7.5}$$

where the φ_n are definite functions, which we suppose normalized, and only the constants a_n depend on $\Psi(q)$. In the integral relation (7.4) this corresponds to a kernel $K_n(q, q')$ which is a product of a function of q and a function of q':

$$K_n(q, q') = \varphi_n(q)\Psi_n^*(q') \tag{7.6}$$

then the linear relation between the constants an and the function $\Psi(q)$ is

$$a_n = \int \Psi(q)\Psi_n^*(q)\mathrm{d}q \tag{7.7}$$

where the $\Psi_n(q)$ are certain functions depending on the process of measurement.

The functions $\varphi_n(q)$ are the normalized wave functions of the electron after measurement. Thus we see how the mathematical formalism of the theory reflects the possibility of finding by measurement a state of the electron described by a definite wave function.

If the measurement is made on an electron with a given wave function $\Psi(q)$, the constants a_n have a simple physical meaning: in accordance with the usual rules, $|a_n|^2$ is the probability that the measurement will give the nth result. The sum of the probabilities of all results is equal to unity:

$$\sum_n |a_n|^2 = 1 \tag{7.8}$$

In order that equations (7.7) and (7.8) should hold for an arbitrary normalized function $\Psi(q)$, it is necessary (cf. §3) that an arbitrary function $\Psi(q)$ can be expanded in terms of the functions $\Psi_n(q)$. This means that the functions $\Psi_n(q)$ form a complete set of normalized and orthogonal functions.

If the initial wave function of the electron coincides with one of the functions $\Psi_n(q)$, then the corresponding constant a_n is evidently equal to unity, while all the others are zero. In other words, a measurement made on an electron in the state $\Psi_n(q)$ gives with certainty the nth result.

All these properties of the functions $\Psi_n(q)$ show that they are the eigenfunctions of some physical quantity (denoted by f) which characterizes the electron, and the measurement concerned can be spoken of as a measurement of this quantity.

It is very important to notice that the functions $\Psi_n(q)$ do not, in general, coincide with the functions $\varphi_n(q)$; the latter are in general not even mutually orthogonal, and do not form a set of eigenfunctions of any operator. This expresses the fact that the results of measurements in quantum mechanics cannot be reproduced. If the electron was in a state $\Psi_n(q)$, then a measurement of the quantity f carried out on it leads with certainty to the value f_n.

After the measurement, however, the electron is in a state $\varphi_n(q)$ different from its initial one, and in this state the quantity f does not in general take any definite value. Hence, on carrying out a second measurement on the electron immediately after the first, we should obtain for f a value which did not agree with that obtained from the first measurement.[18] To predict (in the sense of calculating probabilities) the result of the second measurement from the known result of the first, we must take from the first measurement the wave function $\varphi_n(q)$ of the state in which it resulted, and from the second measurement the wave function $\Psi_n(q)$ of the state whose probability is required. This means that from the equations of quantum mechanics we determine the wave function $\varphi_n(q,t)$ which, at the instant when the first measurement is made, is equal to $\varphi_n(q)$; the probability of the mth result of the second measurement, made at time t, is then given by the squared modulus of the integral $\int \varphi_n(q,t)\Psi_m^*(q)dq$.

We see that the measuring process in quantum mechanics has a "two-faced" character: it plays different parts with respect to the past and future of the electron. With respect to the past, it "verifies" the probabilities of the various possible results predicted from the state brought about by the previous measurement. With respect to the future, it brings about a new state (see also §44). Thus the very nature of the process of measurement involves a far-reaching principle of irreversibility.

This irreversibility is of fundamental significance. We shall see later (at the end of §18) that the basic equations of quantum mechanics are in themselves symmetrical with respect to a change in the sign of the time; here quantum mechanics does not differ from classical mechanics. The irreversibility of the process of measurement, however, causes the two directions of time to be physically non-equivalent, i.e. creates a difference between the future and the past.

---

[18] There is, however, an important exception to the statement that results of measurements cannot be reproduced: the one quantity the result of whose measurement can be exactly reproduced is the coordinate. Two measurements of the coordinates of an electron, made at a sufficiently small interval of time, must give neighbouring values; if this were not so, it would mean that the electron had an infinite velocity. Mathematically, this is related to the fact that the coordinate commutes with the operator of the interaction energy between the electron and the apparatus, since this energy is (in non-relativistic theory) a function of the coordinates only.

CHAPTER II

ENERGY AND MOMENTUM

§ 8. The Hamiltonian operator

The wave function Ψ completely determines the state of a physical system in quantum mechanics. This means that, if this function is given at some instant, not only are all the properties of the system at that instant described, but its behaviour at all subsequent instants is determined (only, of course, to the degree of completeness which is generally admissible in quantum mechanics). The mathematical expression of this fact is that the value of the derivative $\partial\Psi/\partial t$ of the wave function with respect to time at any given instant must be determined by the value of the function itself at that instant, and, by the principle of superposition, the relation between them must be linear. In the most general form we can write

$$i\hbar\frac{\partial\Psi}{\partial t} = \widehat{H}\Psi \tag{8.1}$$

where $\widehat{H}$ is some linear operator; the factor $i\hbar$ is introduced here for a reason that will become apparent.

Since the integral $\int\Psi^*\Psi dq$ is a constant independent of time, we have

$$\frac{d}{dt}\int |\Psi|^2 dq = \int \frac{\partial\Psi^*}{\partial t}\Psi dq + \int \Psi^*\frac{\partial\Psi}{\partial t}dq = 0$$

Substituting here (8.1) and using in the first integral the definition of the transpose of an operator, we can write (omitting the common factor $i/\hbar$)

$$\int \Psi\widehat{H}^*\Psi^* dq - \int \Psi^*\widehat{H}\Psi dq = \int \Psi^*\widetilde{\widehat{H}}^*\Psi dq - \int \Psi^*\widehat{H}\Psi dq$$

$$= \int \Psi^*\left(\widetilde{\widehat{H}}^* - \widehat{H}\right)\Psi dq = 0.$$

Since this equation must hold for an arbitrary function Ψ, it follows that we must have identically $\widehat{H}^\dagger = \widehat{H}$; the operator $\widehat{H}$ is therefore Hermitian. Let

us find the physical quantity to which it corresponds. To do this, we use the limiting expression (6.1) for the wave function and write

$$\frac{\partial \Psi}{\partial t} = \frac{\mathrm{i}}{\hbar} \frac{\partial S}{\partial t} \Psi;$$

the slowly varying amplitude a need not be differentiated. Comparing this equation with the definition (8.1), we see that, in the limiting case, the operator $\widehat{H}$ reduces to simply multiplying by $-\partial S/\partial t$. This means that $-\partial S/\partial t$ is the physical quantity into which the Hermitian operator $\widehat{H}$ passes.

The derivative $-\partial S/\partial t$ is just Hamilton's function H for a mechanical system. Thus the operator $\widehat{H}$ is what corresponds in quantum mechanics to Hamilton's function; this operator is called the *Hamiltonian operator* or, more briefly, the *Hamiltonian* of the system. If the form of the Hamiltonian is known, equation (8.1) determines the wave functions of the physical system concerned. This fundamental equation of quantum mechanics is called the *wave equation.*

§ 9. The differentiation of operators with respect to time

The concept of the derivative of a physical quantity with respect to time cannot be defined in quantum mechanics in the same way as in classical mechanics. For the definition of the derivative in classical mechanics involves the consideration of the values of the quantity at two neighbouring but distinct instants of time. In quantum mechanics, however, a quantity which at some instant has a definite value does not in general have definite values at subsequent instants; this was discussed in detail in §1.

Hence the idea of the derivative with respect to time must be differently defined in quantum mechanics. It is natural to define the *derivative* $\dot{f}$ of a quantity f as the quantity whose mean value is equal to the derivative, with respect to time, of the mean value . Thus we have the definition

$$\dot{\overline{f}} = \overline{\dot{f}} \tag{9.1}$$

Starting from this definition, it is easy to obtain an expression for the quantum-mechanical operator $\widehat{\dot{f}}$ corresponding to the quantity $\widehat{\dot{f}}$:

$$\dot{\overline{f}} = \overline{\dot{f}} = \frac{\mathrm{d}}{\mathrm{d}t} \int \Psi^* \widehat{f} \Psi \mathrm{d}q = \int \Psi^* \frac{\partial \widehat{f}}{\partial t} \Psi \mathrm{d}q + \int \frac{\partial \Psi^*}{\partial t} \widehat{f} \Psi \mathrm{d}q + \int \Psi^* \widehat{f} \frac{\partial \Psi}{\partial t} \mathrm{d}q.$$

Here $\partial f/\partial t$ is the operator obtained by differentiating the operator f with respect to time; $\widehat{f}$ may depend on the time as a parameter. Substituting for $\partial\Psi/\partial t$, $\partial\Psi^*/\partial t$ their expressions according to (8.1), we obtain

$$\overline{\overline{f}} = \int \Psi^* \frac{\partial\widehat{f}}{\partial t}\Psi\mathrm{d}q + \frac{\mathrm{i}}{\hbar}\int(\widehat{H}^*\Psi^*)\widehat{f}\Psi\mathrm{d}q - \frac{\mathrm{i}}{\hbar}\int\Psi^*\widehat{f}(\widehat{H}\Psi)\mathrm{d}q.$$

Since the operator $\widehat{H}$ is Hermitian, we have

$$\int(\widehat{H}^*\Psi^*)(\widehat{f}\Psi)\mathrm{d}q = \int\Psi^*\widehat{H}\widehat{f}\Psi\mathrm{d}q;$$

thus

$$\overline{\overline{f}} = \int\Psi^*\left(\frac{\partial\widehat{f}}{\partial t} + \frac{\mathrm{i}}{\hbar}\widehat{H}\widehat{f} - \frac{\mathrm{i}}{\hbar}\widehat{f}\widehat{H}\right)\Psi\mathrm{d}q.$$

Since, on the other hand, we must have, by the definition of mean values, $\overline{\overline{f}} = \int\Psi^*\widehat{\dot{f}}\Psi\mathrm{d}q$, it is seen that the expression in parentheses in the integrand is the required operator $\widehat{\dot{f}}$:[1]

$$\widehat{\dot{f}} = \frac{\partial f}{\partial t} + \frac{\mathrm{i}}{\hbar}(\widehat{H}\widehat{f} - \widehat{f}\widehat{H}) \qquad (9.2)$$

---

[1] In classical mechanics we have for the total derivative, with respect to time, of a quantity f which is a function of the generalized coordinates q_i and momenta p_i of the system

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_i\left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i\right).$$

Substituting, in accordance with Hamilton's equations, $q_i = \frac{\partial H}{\partial p_i}$ and $p_i = -\frac{\partial H}{\partial q_i}$, we obtain

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [H, f],$$

where

$$[H, f] \equiv \sum_i\left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}\right)$$

is what is called the *Poisson bracket* for the quantities f and H (see Mechanics, §42). On comparing with the expression (9.2), we see that, as we pass to the limit of classical mechanics, the operator $\mathrm{i}(\widehat{H}\widehat{f} - \widehat{f}\widehat{H})$ reduces in the first approximation to zero, as it should, and in the second approximation (with respect to $\hbar$) to the quantity $\hbar\,[H, f]$. This result is true also for any two quantities f and g; the operator $\mathrm{i}(\widehat{f}\widehat{g} - \widehat{g}\widehat{f})$ tends in the limit to the quantity $\hbar\,[f, g]$, where $[f, g]$ is the Poisson bracket

$$[f, g] \equiv \sum_i\left(\frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i} - \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q_i}\right)$$

This follows from the fact that we can always formally imagine a system whose Hamiltonian is $\widehat{g}$.

If the operator $\widehat{f}$ is independent of time, $\dot{\widehat{f}}$ reduces, apart from a constant factor, to the commutator of the operator $\widehat{f}$ and the Hamiltonian.

A very important class of physical quantities is formed by those whose operators do not depend explicitly on time, and also commute with the Hamiltonian, so that $\dot{\widehat{f}} = 0$. Such quantities are said to be *conserved*. For these $\overline{\dot{f}} = \dot{\overline{f}} = 0$, that is, is constant. In other words, the mean value of the quantity remains constant in time. We can also assert that, if in a given state the quantity f has a definite value (i.e. the wave function is an eigenfunction of the operator $\widehat{f}$), then it will have a definite value (the same one) at subsequent instants also.

§ 10. Stationary states

The Hamiltonian of a closed system (and of a system in a *constant* external field) cannot contain the time explicitly. This follows from the fact that, for such a system, all times are equivalent. Since, on the other hand, any operator of course commutes with itself, we reach the conclusion that Hamilton's function is conserved for systems which are not in a varying external field. As is well known, a Hamilton's function which is conserved is called the *energy*. The law of conservation of energy in quantum mechanics signifies that, if in a given state the energy has a definite value, this value remains constant in time.

States in which the energy has definite values are called *stationary states* of a system. They are described by wave functions Ψ_n which are the eigenfunctions of the Hamiltonian operator, i.e. which satisfy the equation $\widehat{H}\Psi_n = E_n\Psi_n$, where E_n are the eigenvalues of the energy. Correspondingly, the wave equation (8.1) for the function Ψ_n,

$$\mathrm{i}\hbar\frac{\partial\Psi_n}{\partial t} = \widehat{H}\Psi_n = E_n\Psi_n$$

can be integrated at once with respect to time and gives

$$\Psi_n = \exp\left(-\frac{\mathrm{i}}{\hbar}E_n t\right)\psi_n(q) \tag{10.1}$$

where Ψ_n is a function of the coordinates only. This determines the relation between the wave functions of stationary states and the time.

We shall denote by the small letter ψ the wave functions of stationary states without the time factor. These functions, and also the eigenvalues of the energy, are determined by the equation

$$\widehat{H}\psi = E\psi \tag{10.2}$$

The stationary state with the smallest possible value of the energy is called the *normal* or *ground state* of the system.

The expansion of an arbitrary wave function Ψ in terms of the wave functions of stationary states has the form

$$\Psi = \sum_n a_n \exp\left(-\frac{i}{\hbar}E_n t\right)\psi_n(q) \tag{10.3}$$

The squared moduli $|a_n|^2$ of the expansion coefficients, as usual, determine the probabilities of various values of the energy of the system.

The probability distribution for the coordinates in a stationary state is determined by the squared modulus $|\Psi_n|^2 = |\psi_n|2$ we see that it is independent of time. The same is true of the mean values

$$\overline{f} = \int \Psi_n^* \widehat{f}\Psi_n \mathrm{d}q = \int \psi_n^* \widehat{f}\psi_n \mathrm{d}q$$

of any physical quantity f (whose operator does not depend explicitly on the time).

As has been said, the operator of any quantity that is conserved commutes with the Hamiltonian. This means that any physical quantity that is conserved can be measured simultaneously with the energy.

Among the various stationary states, there may be some which correspond to the same value of the energy (the same *energy level* of the system), but differ in the values of some other physical quantities. Such energy levels, to which several different stationary states correspond, are said to be *degenerate*. Physically, the possibility that degenerate levels can exist is related to the fact that the energy does not in general form by itself a complete set of physical quantities.

If there are two conserved physical quantities f and g whose operators do not commute, then the energy levels of the system are in general degenerate. For, let ψ be the wave function of a stationary state in which, besides the energy, the quantity f also has a definite value. Then we can say that the function $\widehat{g}\psi$ does not coincide (apart from a constant factor) with ψ; if it did, this would mean that the quantity g also had a definite value, which is impossible, since f and g cannot be measured simultaneously. On the other hand, the function $\widehat{g}\psi$ is an eigenfunction of the Hamiltonian, corresponding to the same value E of the energy as ψ:

$$\widehat{H}(\widehat{g}\psi) = \widehat{g}\widehat{H}\psi = E(\widehat{g}\psi)$$

Thus we see that the energy E corresponds to more than one eigenfunction, i.e. the energy level is degenerate.

It is clear that any linear combination of wave functions corresponding to the same degenerate energy level is also an eigenfunction for that value of the energy. In other words, the choice of eigenfunctions of a degenerate energy level is not unique. Arbitrarily selected eigenfunctions of a degenerate energy level are not, in general, orthogonal. By a proper choice of linear combinations of them, however, we can always obtain a set of orthogonal (and normalized) eigenfunctions (and this can be done in infinitely many ways; for the number of independent coefficients in a linear transformation of n functions is n^2, while the number of normalization and orthogonality conditions for n functions is $n(n + 1)/2$, i.e. less than n^2).

These statements concerning the eigenfunctions of a degenerate energy level relate, of course, not only to eigenfunctions of the energy, but also to those of any operator. Only those functions are automatically orthogonal which correspond to different eigenvalues of the operator concerned; functions which correspond to the same degenerate eigenvalue are not in general orthogonal.

If the Hamiltonian of the system is the sum of two (or more) parts, $\widehat{H} = \widehat{H}_1 + \widehat{H}_2$, one of which contains only the coordinates q_1 and the other only the coordinates q_2, then the eigenfunctions of the operator $\widehat{H}$ can be written down as products of the eigenfunctions of the operators $\widehat{H}_1$ and $\widehat{H}_2$, and the eigenvalues of the energy are equal to the sums of the eigenvalues of these operators.

The spectrum of eigenvalues of the energy may be either discrete or continuous. A stationary state of a discrete spectrum always corresponds to a finite motion of the system, i.e. one in which neither the system nor any part of it moves off to infinity. For, with eigenfunctions of a discrete spectrum, the integral $\int |\Psi|^2 \mathrm{d}q$, taken over all space, is finite. This certainly means that the squared modulus $|\Psi|^2$ decreases quite rapidly, becoming zero at infinity. In other words, the probability of infinite values of the coordinates is zero; that is, the system executes a finite motion, and is said to be in a *bound state*.

For wave functions of a continuous spectrum, the integral $\int |\Psi|^2 \mathrm{d}q$ diverges. Here the squared modulus $|\Psi|^2$ of the wave function does not directly determine the probability of the various values of the coordinates, and must be regarded only as a quantity proportional to this probability. The divergence of the integral $\int |\Psi|^2 \mathrm{d}q$ is always due to the fact that $|\Psi|^2$ does not become zero at infinity (or becomes zero insufficiently rapidly). Hence we can say that the integral $\int |\Psi|^2 \mathrm{d}q$, taken over the region of space outside any arbitrarily large but finite closed surface, will always diverge. This means that, in the state considered, the system (or some part of it) is at infinity.

For a wave function which is a superposition of the wave functions of various stationary states of a continuous spectrum, the integral $\int |\Psi|^2 \mathrm{d}q$ may converge, so that the system lies in a finite region of space. However, in the course of time, this region moves unrestrictedly, and eventually the system moves off to infinity. This can be seen as follows. Any superposition of wave functions of a continuous spectrum has the form

$$\Psi = \int a_E \exp\left(-\frac{\mathrm{i}}{\hbar}Et\right)\psi_E(q)\mathrm{d}E$$

The squared modulus of Ψ can be written in the form of a double integral:

$$|\Psi|^2 = \iint a_E a_{E'}^* \exp\left(\frac{\mathrm{i}}{\hbar}(E' - E)t\right)\psi_E(q)\psi_{E'}^*(q)\mathrm{d}E\mathrm{d}E'.$$

If we average this expression over some time interval T, and then let T tend to infinity, the mean values of the oscillating factors $\exp\{\mathrm{i}(E' - E)t/\hbar\}$, and therefore the whole integral, tend to zero in the limit. Thus the mean value, with respect to time, of the probability of finding the system at any given point of configuration space tends to zero. This is possible only if the motion takes place throughout infinite space.[2] Thus the stationary states of a continuous spectrum correspond to an infinite motion of the system.

§ 11. Matrices

We shall suppose for convenience that the system considered has a discrete energy spectrum; all the relations obtained below can be generalized at once to the case of a continuous spectrum. Let $\Psi = \sum a_n \Psi_n$ be the expansion of an arbitrary wave function in terms of the wave functions Ψ_n of the stationary states. If we substitute this expansion in the definition (3.8) of the mean value of some quantity f, we obtain

$$\overline{f} = \sum_n \sum_m a_n^* a_m f_{nm}(t) \tag{11.1}$$

---

[2]) Note that, for a function which is a superposition of functions of a discrete spectrum, we should have

$$\overline{|\Psi|^2} = \sum_n \sum_m a_n a_m^* \overline{\exp\left\{\frac{\mathrm{i}}{\hbar}(E_m - E_n)t\right\}}\psi_n\psi_m^* = \sum_n |a_n\psi_n(q)|^2,$$

i.e. the probability density remains finite on averaging over time.

where $f_{nm}(t)$ denotes the integral

$$f_{nm}(t) = \int \Psi_n^* \widehat{f} \Psi_m \mathrm{d}q \tag{11.2}$$

The set of quantities $f_{nm}(t)$ with all possible n and m is called the *matrix* of the quantity f, and each of the $f_{nm}(t)$ is called the *matrix element* corresponding to the *transition* from state m to state n.[3]

The dependence of the matrix elements $f_{nm}(t)$ on time is determined (if the operator does not contain the time explicitly) by the dependence of the functions Ψ_n on time. Substituting for them the expressions (10.1), we find that

$$f_{nm}(t) = f_{nm}\mathrm{e}^{\mathrm{i}\omega_{nm}t}, \tag{11.3}$$

where

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} \tag{11.4}$$

is what is called the *transition frequency* between the states n and m, and the quantities

$$f_{nm} = \int \psi_n^* \widehat{f} \psi_m \mathrm{d}q \tag{11.5}$$

form the matrix of the quantity f which is independent of time, and which is commonly used.[4]

The matrix elements of the derivative are obtained by differentiating the matrix elements of the quantity f with respect to time; this follows directly from the fact that

$$\overline{\dot{f}} = \dot{\overline{f}} = \sum_m \sum_n a_n^* a_m \dot{f}_{nm}(t). \tag{11.6}$$

From (11.3) we thus have for the matrix elements of $\dot{f}$

$$\dot{f}_{nm}(t) = \mathrm{i}\omega_{nm} f_{nm}(t) \tag{11.7}$$

or (cancelling the time factor $\exp(\mathrm{i}\omega_{nm}t)$ from both sides) for the matrix elements independent of time

$$\left(\dot{f}\right)_{nm} = \mathrm{i}\omega_{nm} f_{nm} = \frac{\mathrm{i}}{\hbar}(E_n - E_m)f_{nm} \tag{11.8}$$

---

[3] The matrix representation of physical quantities was introduced by Heisenberg in 1925, before Schrödinger's discovery of the wave equation. "Matrix mechanics" was later developed by M. Born, W. Heisenberg and P. Jordan.

[4] Because of the indeterminacy of the phase factor in normalized wave functions (see §2), the matrix elements f_{nm} (and $f_{nm}(t)$) also are determined only to within a factor of the form $\exp\left[\mathrm{i}(\alpha_m - \alpha_n)\right]$. Here again this indeterminacy has no effect on any physical results.

To simplify the notation in the formulae, we shall derive all our relations below for the matrix elements independent of time; exactly similar relations hold for the matrices which depend on the time.

For the matrix elements of the complex conjugate f^* of the quantity f we obtain, taking into account the definition of the Hermitian conjugate operator,

$$(f^*)_{nm} = \int \psi_n^* \widehat{f}^\dagger \psi_m \mathrm{d}q = \int \psi_n^* \widetilde{\widehat{f}^*} \psi_m \mathrm{d}q = \int \psi_m \widehat{f}^* \psi_n^* \mathrm{d}q$$

or

$$(f^*)_{nm} = (f_{mn})^* \tag{11.9}$$

For real physical quantities, which are the only ones we usually consider, we consequently have

$$f_{nm} = f_{mn}^* \tag{11.10}$$

(f_{mn}^* stands for $(f_{mn})^*$). Such matrices, like the corresponding operators, are said to be *Hermitian.*

Matrix elements with $n = m$ are called diagonal elements. These are independent of time, and (11.10) shows that they are real. The element f_{nn} is the mean value of the quantity f in the state Ψ_n.

It is not difficult to obtain the multiplication rule" for matrices. To do so, we first observe that the formula

$$\widehat{f}\psi_n = \sum_m f_{mn}\psi_m \tag{11.11}$$

holds. This is simply the expansion of the function in terms of the functions ψ_m, the coefficients being determined in accordance with the general formula (3.5). Remembering this formula, let us write down the result of the product of two operators acting on the function ψ_n:

$$\widehat{f}\widehat{g}\psi_n = \widehat{f}(\widehat{g}\psi_n) = \widehat{f}\sum_k g_{kn}\psi_k = \sum_k g_{kn}\widehat{f}\psi_k = \sum_{k,m} g_{kn}f_{mk}\psi_m.$$

Since, on the other hand, we must have

$$\widehat{f}\widehat{g}\psi_n = \sum_m (fg)_{mn}\psi_m,$$

we arrive at the result that the matrix elements of the product fg are determined by the formula

$$(fg)_{mn} = \sum_k f_{mk}g_{kn} \tag{11.12}$$

This rule is the same as that used in mathematics for the multiplication of matrices: the rows of the first matrix in the product are multiplied by the columns of the second matrix.

If the matrix is given, then so is the operator itself. In particular, if the matrix is given, it is in principle possible to determine the eigenvalues of the physical quantity concerned and the corresponding eigenfunctions.

We shall now consider the values of all quantities at some definite instant, and expand an arbitrary wave function Ψ (at that instant) in terms of the eigenfunctions of the Hamiltonian, i.e. of the wave functions ψ_m of the stationary states (these wave functions are independent of time).

$$\Psi = \sum_m c_m \psi_m \tag{11.13}$$

where the expansion coefficients are denoted by c_m. We substitute this expansion in the equation $\widehat{f}\Psi = f\Psi$ which determines the eigenvalues and eigenfunctions of the quantity f. We have

$$\sum_m c_m(\widehat{f}\psi_m) = f \sum_m c_m \psi_m$$

We multiply both sides of this equation by ψ_n^* and integrate over q. Each of the integrals on the left-hand side of the equation is the corresponding matrix element f_{nm}. On the right-hand side, all the integrals $\int \psi_n^* \psi_m \mathrm{d}q$ with $m \neq n$ vanish by virtue of the orthogonality of the functions ψ_m, and $\int \psi_n^* \psi_n \mathrm{d}q = 1$ by virtue of their normalization.[5] Thus

$$\sum_m f_{nm} c_m = f c_n, \tag{11.14}$$

or

$$\sum_m (f_{nm} - f\delta_{nm})c_m = 0,$$

where

$$\delta_{nm} = \begin{cases} 0, & n \neq m; \\ 1, & n = m. \end{cases}$$

---

[5] In accordance with the general rule (§5), the set of coefficients c_n in the expansion (11.13) can be considered as the wave function in the "energy representation" (the variable being the suffix n that gives the number of the energy eigenvalue). The matrix f_{nm} here acts as the operator in this representation, the action of which on the wave function is given by the left-hand side of (11.14). The formula then corresponds to the general expression for the mean value of a quantity in terms of its operator and the wave function of the state concerned.

Thus we have obtained a system of homogeneous algebraic equations of the first degree (with the c_m as unknowns). As is well known, such a system has solutions different from zero only if the determinant formed by the coefficients in the equations vanishes, i.e. only if

$$|f_{nm} - f\delta_{nm}| = 0 \tag{11.15}$$

The roots of this equation (in which f is regarded as the unknown) are the possible values of the quantity f. The set of values c_m satisfying the equations (11.14) when f is equal to any of these values determines the corresponding eigenfunction.

If, in the definition (11.5) of the matrix elements of the quantity f, we take as ψ_n the eigenfunctions of this quantity, then from the equation $\widehat{f}\psi_n = f_n\psi_n$ we have

$$f_{nm} = \int \psi_n^*\widehat{f}\psi_m \mathrm{d}q = f_m \int \psi_n^*\psi_m \mathrm{d}q.$$

By virtue of the orthogonality and normalization of the functions ψ_m, this gives $f_{nm} = 0$ for $n \neq m$ and $f_{mm} = f_m$. Thus only the diagonal matrix elements are different from zero, and each of these is equal to the corresponding eigenvalue of the quantity f. A matrix with only these elements different from zero is said to be put in *diagonal form*. In particular, in the usual representation, with the wave functions of the stationary states as the functions ψ_n, the energy matrix is diagonal (and so are the matrices of all other physical quantities having definite values in the stationary states). In general, the matrix of a quantity f, defined with respect to the eigenfunctions of some operator $\widehat{g}$, is said to be the matrix of f *in a representation in which g is diagonal*. We shall always, except where the subject is specially mentioned, understand in future by the matrix of a physical quantity its matrix in the usual representation, in which the energy is diagonal. Everything that has been said above regarding the dependence of matrices on time refers, of course, only to this usual representation.[6]

By means of the matrix representation of operators we can prove the theorem mentioned in §4: if two operators commute with each other, they have their entire sets of eigenfunctions in common. Let $\widehat{f}$ and $\widehat{g}$ be two such operators. From $\widehat{f}\widehat{g} = \widehat{g}\widehat{f}$ and the matrix multiplication rule (11.12), it follows that

$$\sum_k f_{mk}g_{kn} = \sum_k g_{mk}f_{kn}.$$

---

[6] Bearing in mind the diagonality of the energy matrix, it is easy to see that equation (11.8) is the operator relation (9.2) written in matrix form.

If we take the eigenfunctions of the operator $\widehat{f}$ as the set of functions ψ_n with respect to which the matrix elements are calculated, we shall have $f_{mk} = 0$ for $m \neq k$, so that the above equation reduces to $f_{mm}g_{mn} = g_{mn}f_{nn}$, or

$$g_{mn}(f_m - f_n) = 0$$

If all the eigenvalues f_n of the quantity f are different, then for all $m \neq n$ we have $f_m f_n \neq 0$, so that we must have $g_{mn} = 0$. Thus the matrix g_{mn} is also diagonal, i.e. the functions ψ_n are eigenfunctions of the physical quantity g also. If, among the values f_n, there are some which are equal (i.e. if there are eigenvalues to which several different eigenfunctions correspond), then the matrix elements g_{mn} corresponding to each such group of functions ψ_n are, in general, different from zero. However, linear combinations of the functions ψ_n which correspond to a single eigenvalue of the quantity f are evidently also eigenfunctions of f; one can always choose these combinations in such a way that the corresponding non-diagonal matrix elements g_{mn} are zero, and thus, in this case also, we obtain a set of functions which are simultaneously the eigenfunctions of the operators $\widehat{f}$ and $\widehat{g}$.

The following formula is useful in applications:

$$\left(\frac{\partial H}{\partial \lambda}\right)_{nn} = \frac{\partial E_n}{\partial \lambda} \tag{11.16}$$

where λ is a parameter on which the Hamiltonian $\widehat{H}$ (and therefore the energy eigenvalues E_n) depends. It is proved as follows. Differentiating the equation $(\widehat{H} - E_n)\psi_n = 0$ with respect to λ and then multiplying on the left by ψ_n^*, we obtain

$$\psi_n^*(\widehat{H} - E_n)\frac{\partial \psi_n}{\partial \lambda} = \psi_n^*\left(\frac{\partial E_n}{\partial \lambda} - \frac{\partial \widehat{H}}{\partial \lambda}\right)\psi_n$$

On integration with respect to q, the left-hand side gives zero, since

$$\int \psi_n^*(\widehat{H} - E_n)\frac{\partial \psi_n}{\partial \lambda}\mathrm{d}q = \int \frac{\partial \psi_n}{\partial \lambda}(\widehat{H} - E_n)^*\psi_n^*\mathrm{d}q,$$

the operator $\widehat{H}$ being Hermitian. The right-hand side gives the required equation.

A widely used notation (introduced by Dirac) in recent literature is that which denotes the matrix elements f_{nm} by[7]

$$\langle n|f|m\rangle \tag{11.17}$$

---

[7] Both notations are used in the present book. The form (11.17) is especially convenient when each suffix has to be written as several letters.

This symbol is written so that it may be regarded as "consisting" of the quantity f and the symbols $|m\rangle$ and $\langle n|$ which respectively stand for the initial and final states as such (independently of the representation of the wave functions of the states). With the same symbols we can construct notations for the expansion coefficients of wave functions: if there is a complete set of wave functions corresponding to the states $|n_1\rangle, |n_2\rangle, \ldots$, the coefficients in the expansion in terms of these of the wave function of a state $|m\rangle$ are denoted by

$$\langle n_i|m\rangle = \int \psi_{n_i}^* \psi_m \mathrm{d}q. \tag{11.18}$$

§ 12. Transformation of matrices

The matrix elements of a given physical quantity can be defined with respect to various sets of wave functions, for example the wave functions of stationary states described by various sets of physical quantities, or the wave functions of stationary states of the same system in various external fields. The problem therefore arises of the transformation of matrices from one representation to another.

Let $\psi_n(q)$ and $\psi'_n(q)$ $(n = 1, 2, \ldots)$ be two complete sets of orthonormal functions, related by some linear transformation:

$$\psi'_n = \sum_m S_{mn} \psi_m \tag{12.1}$$

which is simply an expansion of the function ψ'_n in terms of the complete set of functions ψ_n. This transformation may be conventionally written in the operator form

$$\psi'_n = \widehat{S} \psi_n \tag{12.2}$$

The operator $\widehat{S}$ must satisfy a certain condition in order that the functions ψ'_n should be orthonormal if the functions ψ_n are. Substituting (12.2) in the condition

$$\int \psi_m'^* \psi_n' \mathrm{d}q = \delta_{nm}$$

and using the definition of the transposed operator (3.14), we have

$$\int (\widehat{S}\psi_n)\widehat{S}^*\psi_m^* \mathrm{d}q = \int \psi_m^* \widetilde{\widehat{S}^*} \widehat{S} \psi_n \mathrm{d}q = \delta_{nm}$$

If these equations hold for all m and n, we must have $\widetilde{\widehat{S}^*}\widehat{S} = 1$, or

$$\widetilde{\widehat{S}^*} \equiv \widehat{S}^\dagger = \widehat{S}^{-1} \tag{12.3}$$

i.e. the inverse operator is equal to the Hermitian conjugate operator. Operators having this property are said to be unitary. Owing to this property, the transformation $\psi_n = \widehat{S}^{-1}\psi'_n$ inverse to (12.1) is given by

$$\psi_n = \sum_m S^*_{nm}\psi'_m \tag{12.4}$$

Writing the equations $\widehat{S}^\dagger\widehat{S} = 1$ and $\widehat{S}\widehat{S}^\dagger = 1$ in matrix form, we obtain the following forms of the *unitarity* condition:

$$\sum_l S^*_{lm}S_{ln} = \delta_{mn} \tag{12.5}$$

$$\sum_l S^*_{ml}S_{nl} = \delta_{mn} \tag{12.6}$$

Let us now consider some physical quantity f and write down its matrix elements in the "new" representation, i.e. with respect to the functions ψ'_n. These are given by the integrals

$$\int \psi'^*_m \widehat{f}\psi'_n \mathrm{d}q = \int (\widehat{S}^*\psi^*_m)(\widehat{f}\widehat{S}\psi_n)\mathrm{d}q$$
$$= \int \psi^*_m \widetilde{\widehat{S}^*}\widehat{f}\widehat{S}\psi_n \mathrm{d}q = \int \psi^*_m \widehat{S}^{-1}\widehat{f}\widehat{S}\psi_n \mathrm{d}q.$$

Hence we see that the matrix of the operator in the new representation is equal to the matrix of the operator

$$\widehat{f}' = \widehat{S}^{-1}\widehat{f}\widehat{S} \tag{12.7}$$

in the old representation.[8])

The sum of the diagonal elements of a matrix is called the *trace* or *spur*[9]) of the matrix and denoted by $\mathrm{tr}f$:

$$\mathrm{tr}f = \sum_n f_{nn} \tag{12.8}$$

---

[8]) If $\{\widehat{f},\widehat{g}\} = -\mathrm{i}\hbar\widehat{c}$ is the commutation rule for two operators $\widehat{f}$ and $\widehat{g}$, the transformation (12.7) gives $\{\widehat{f}',\widehat{g}'\} = -\mathrm{i}\hbar\widehat{c}'$, i.e. the rule is unchanged. We have shown in the footnote in §9 that is the quantum analogue of the classical Poisson bracket $[f, g]$. In classical mechanics, however, the Poisson brackets are invariant under canonical transformations of the variables (generalized coordinates and momenta); see Mechanics, §45. In this sense we can say that unitary transformations in quantum mechanics play a role analogous to that of canonical transformations in classical mechanics.

[9]) From the German word *Spur*. The notation $\mathrm{sp}f$ is also used. The trace can be defined, of course, only if the sum over n is convergent.

It may be noted first of all that the trace of a product of two matrices is independent of the order of multiplication:

$$\text{tr}(fg) = \text{tr}(gf) \tag{12.9}$$

since the rule of matrix multiplication gives

$$\text{tr}(fg) = \sum_n \sum_n f_{nk} g_{kn} = \sum_k \sum_n g_{kn} f_{nk} = \text{tr}(gf).$$

Similarly we can easily see that, for a product of several matrices, the trace is unaffected by a cyclic permutation of the factors; for example,

$$\text{tr}(fgh) = \text{tr}(hfg) = \text{tr}(ghf) \tag{12.10}$$

An important property of the trace is that it does not depend on the choice of the set of functions with respect to which the-matrix elements are defined, since

$$(\text{tr}f)' = \text{tr}(S^{-1}fS) = \text{tr}(S^{-1}Sf) = \text{tr}f \tag{12.11}$$

A unitary transformation leaves unchanged the sum of the squared moduli of the functions that are transformed: from (12.6) we have

$$\sum_i |\psi_i'|^2 = \sum_{k,l,i} S_{ki}\psi_k S_{li}^* \psi_l^* = \sum_{k,l} \psi_k \psi_l^* \delta_{kl} = \sum_k |\psi_k|^2. \tag{12.12}$$

Any unitary operator may be written-as

$$\widehat{S} = e^{i\widehat{R}} \tag{12.13}$$

where $\widehat{R}$ is an Hermitian operator: since $\widehat{R}^\dagger = \widehat{R}$, we have

$$\widehat{S}^\dagger = e^{-i\widehat{R}^\dagger} = e^{-i\widehat{R}} = \widehat{S}^{-1}$$

The expansion

$$\widehat{f}' = \widehat{S}^{-1}\widehat{f}\widehat{S} = \widehat{f} + \{\widehat{f}, i\widehat{R}\} + \frac{1}{2}\{\{\widehat{f}, i\widehat{R}\}, i\widehat{R}\} + \dots \tag{12.14}$$

is easily verified by a direct expansion of the factors $\exp(\pm i\widehat{R})$ in powers of $\widehat{R}$. This expansion may be useful when $\widehat{R}$ is proportional to a small parameter, so that (12.14) becomes an expansion in powers of the parameter.

§ 13. The Heisenberg representation of operators

In the mathematical formalism of quantum mechanics described here, the operators corresponding to various physical quantities act on functions of the coordinates and do not usually depend explicitly on time. The time dependence of the mean values of physical quantities is due only to the time dependence of the wave function of the state, according to the formula

$$\overline{f}(t) = \int \Psi^*(q,t)\widehat{f}\Psi(q,t)\mathrm{d}q. \tag{13.1}$$

The quantum-mechanical treatment can, however, be formulated also in a somewhat different but equivalent form, in which the time dependence is transferred from the wave functions to the operators. Although we shall not use this *Heisenberg representation* (as opposed to the *Schrödinger representation*) of operators in the present volume, a statement of it is given here with a view to applications in the relativistic theory.

We define the operator (which is unitary; see (12.13))

$$\widehat{S} = \exp\left(-\frac{\mathrm{i}}{\hbar}\widehat{H}t\right), \tag{13.2}$$

where $\widehat{H}$ is the Hamiltonian of the system. By definition, its eigenfunctions are the same as those of the operator $\widehat{H}$, i.e. the stationary-state wave functions $\psi_n(q)$, where

$$\widehat{S}\psi_n(q) = \exp(-\frac{\mathrm{i}}{\hbar}E_nt)\psi_n(q). \tag{13.3}$$

Hence it follows that the expansion (10.3) of an arbitrary wave function Ψ in terms of the stationary-state wave functions can be written in the operator form

$$\Psi(q,t) = \widehat{S}\Psi(q,0), \tag{13.4}$$

i.e. the effect of the operator $\widehat{S}$ is to convert the wave function of the system at some initial instant into the wave function at an arbitrary instant.

Defining, as in (12.7), the time-dependent operator

$$\widehat{f}(t) = \widehat{S}^{-1}\widehat{f}\widehat{S} \tag{13.5}$$

we have

$$\overline{f}(t) = \int \Psi^*(q,0)\widehat{f}(t)\Psi(q,0)\mathrm{d}q, \tag{13.6}$$

and thus obtain the formula (3.8) for the mean value of the quantity f in a form in which the time dependence is entirely transferred to the operator (for our definition of an operator rests on formula (3.8)).

It is evident that the matrix elements of the operator (13.5) with respect to the stationary-state wave functions are the same as the time-dependent matrix elements $f_{nm}(t)$ defined by formula (11.3).

Finally, differentiating the expression (13.5) with respect to time (assuming that the operators $\widehat{f}$ and $\widehat{H}$ do not themselves involve t), we obtain

$$\frac{\partial}{\partial t}\widehat{f}(t) = \frac{\mathrm{i}}{\hbar}\left[\widehat{H}\widehat{f}(t) - \widehat{f}(t)\widehat{H}\right], \tag{13.7}$$

which is similar in form to (9.2) but has a somewhat different significance: the expression (9.2) defines the operator $\widehat{\dot{f}}$ corresponding to the physical quantity $\dot{f}$, while the left-hand side of equation (13.7) is the time derivative of the operator of the quantity f itself.

§ 14. The density matrix

The description of a system by means of a wave function is the most complete description possible in quantum mechanics, in the sense indicated at the end of §1.

States that do not allow such a description are encountered if we consider a system that is part of a larger closed system. We suppose that the closed system as a whole is in some state described by the wave function $\Psi(q, x)$, where x denotes the set of coordinates of the system considered, and q the remaining coordinates of the closed system. This function in general does not fall into a product of functions of x and of q alone, so that the system does not have its own wave function.[10]

Let f be some physical quantity pertaining to the system considered. Its operator therefore acts only on the coordinates x, and not on q. The mean value of this quantity in the state considered is

$$\overline{f} = \iint \Psi^*(q, x)\widehat{f}\Psi(q, x)\mathrm{d}q\mathrm{d}x. \tag{14.1}$$

---

[10] In order that $\Psi(q, x)$ should (at a given instant) fall into such a product, the measurement as a result of which this state was brought about must completely describe the system considered and the remainder of the closed system separately. In order that $\Psi(q, x)$ should continue to have this form at subsequent instants, it is necessary in addition that these parts of the closed system should not interact (see §2). Neither of these conditions is now assumed.

We introduce the function $\rho(x, x')$ defined by

$$\rho(x, x') = \int \Psi(q, x) \Psi^*(q, x') \mathrm{d}q \tag{14.2}$$

where the integration is extended only over the coordinates q; this function is called the *density matrix* of the system. From the definition (14.2) it is evident that the function is "Hermitian":

$$\rho^*(x, x') = \rho(x', x) \tag{14.3}$$

The "diagonal elements" of the density matrix

$$\rho(x, x) = \int |\Psi(q, x)|^2 \mathrm{d}q$$

determine the probability distribution for the coordinates of the system.

Using the density matrix, we can write the mean value f in the form

$$\overline{f} = \int \left[\widehat{f} \rho(x, x')_{x'=x} \mathrm{d}x \right] \tag{14.4}$$

Here $\widehat{f}$ acts only on the variables x in the function $\rho(x, x')$; after calculating the result of its action, we put $x' = x$. We see that, if we know the density matrix, we can calculate the mean value of any quantity characterizing the system. It follows from this that, by means of $\rho(x, x')$, we can also determine the probabilities of various values of the physical quantities in the system. Thus the state of a system which does not have a wave function can be described by means of a density matrix. This does not contain the coordinates q which do not belong to the system concerned, though, of course, it depends essentially on the state of the closed system as a whole.

The description by means of the density matrix is the most general form of quantum-mechanical description of the system. The description by means of the wave function, on the other hand, is a particular case of this, corresponding to a density matrix of the form $\rho(x, x') = \Psi(x) \Psi^*(x')$. The following important difference exists between this particular case and the general one.[11] For a state having a wave function there is always a complete set of measuring processes such that they lead with certainty to definite results (mathematically, this means that Ψ is an eigenfunction of some operator). For states having only a density matrix, on the other hand, there is no complete set of measuring processes whose result can be uniquely predicted.

---

[11] States having a wave function are called "pure" states, as distinct from "mixed" states, which are described by a density matrix.

Let us now suppose that the system is closed, or became so at some instant. Then we can derive an equation giving the change in the density matrix with time, similar to the wave equation for the Ψ function. The derivation can be simplified by noticing that the required linear differential equation for $\rho(x, x', t)$ must be satisfied in the particular case where the system has a wave function, i.e.

$$\rho(x, x', t) = \Psi(x, t)\Psi^*(x, t).$$

Differentiating with respect to time and using the wave equation (8.1), we have

$$i\hbar\frac{\partial\rho}{\partial t} = i\hbar\Psi^*(x', t)\frac{\partial\Psi(x, t)}{\partial t} + i\hbar\Psi(x, t)\frac{\partial\Psi^*(x', t)}{\partial t} =$$
$$= \Psi^*(x', t)\widehat{H}\Psi(x, t) - \Psi(x, t)\widehat{H}'^*\Psi^*(x', t),$$

where $\widehat{H}$ is the Hamiltonian of the system, acting on a function of x, and $\widehat{H}'$ is the same operator acting on a function of x'. The functions $\Psi^*(x', t)$ and $\Psi(x, t)$ can obviously be placed behind the respective operators $\widehat{H}$ and $\widehat{H}'$, and we thus obtain the required equation:

$$i\hbar\frac{\partial\rho(x, x', t)}{\partial t} = (\widehat{H} - \widehat{H}'^*)\rho(x, x', t). \qquad (14.5)$$

Let $\Psi_n(x, t)$ be the wave functions of the stationary states of the system, i.e. the eigenfunctions of its Hamiltonian. We expand the density matrix in terms of these functions; the expansion consists of a double series in the form

$$\rho(x, x', t) = \sum_m\sum_n a_{mn}\Psi_n^*(x', t)\Psi_m(x, t)$$
$$= \sum_m\sum_n \psi_n^*(x')\psi_m(x)\exp\left(\frac{i}{\hbar}(E_n - E_m)t\right). \qquad (14.6)$$

For the density matrix, this expansion plays a part analogous to that of the expansion (10.3) for wave functions. Instead of the set of coefficients a_n, we have here the double set of coefficients a_{mn}. These clearly have the property of being "Hermitian", like the density matrix itself:

$$a_{nm}^* = a_{mn}. \qquad (14.7)$$

For the mean value of some quantity f we have, substituting (14.6) in (14.4),

$$\overline{f} = \sum_m\sum_n a_{mn}\int \Psi_n^*(x, t)\widehat{f}\Psi_m(x, t)\mathrm{d}x,$$

or

$$\overline{f} = \sum_m \sum_n a_{mn} f_{nm}(t) = \sum_m \sum_n a_{mn} f_{nm} \exp\left(\frac{\mathrm{i}}{\hbar}(E_n - E_m)t\right), \quad (14.8)$$

where f_{nm} are the matrix elements of the quantity f. This expression is similar to formula (11.1).[12]

The quantities a_{mn} must satisfy certain inequalities. The "diagonal elements" $\rho(x, x)$ of the density matrix, which determine the probability distribution for the coordinates, must obviously be positive quantities. It therefore follows from the expression (14.6) (with $x' = x$) that the quadratic form

$$\sum_n \sum_m a_{mn} \xi_n^* \xi_m$$

constructed with the coefficients a_{mn} (where the ξ_n are arbitrary complex quantities) must be positive. This places certain conditions, known from the theory of quadratic forms, on the quantities a_{nm}. In particular, all the "diagonal" quantities must clearly be positive:

$$a_{nm} \geqslant 0 \qquad (14.9)$$

and any three quantities a_{nn}, a_{mm} and a_{mn} must satisfy the inequality

$$a_{nn} a_{mm} \geqslant |a_{mn}|^2 \qquad (14.10)$$

To the "pure" case, where the density matrix reduces to a product of functions, there evidently corresponds a matrix a_{mn} of the form

$$a_{mn} = a_m a_n^*. \qquad (14.11)$$

We shall indicate a simple criterion which enables us to decide, from the form of the matrix a_{mn}, whether we are concerned with a "pure" or a "mixed" state. In the pure case we have

$$(a^2)_{mn} = \sum_k a_{mk} a_{kn} = \sum_k a_k^* a_m a_n^* a_k = a_m a_n^* \sum_k |a_k|^2 = a_m a_n^*,$$

or

$$(a^2)_{mn} = a_{mn}, \qquad (14.12)$$

i.e. the density matrix is equal to its own square.

---

[12] The quantities a_{mn} form the density matrix in the energy representation. The description of the states of a system by means of this matrix was introduced independently by L. Landau and F. Bloch in 1927.

§ 15. Momentum

Let us consider a closed system of particles not in an external field. Since all positions in space of such a system as a whole are equivalent, we can say, in particular, that the Hamiltonian of the system does not vary when the system undergoes a parallel displacement over any distance. It is sufficient that this condition should be fulfilled for an arbitrary small displacement.

An infinitely small parallel displacement over a distance $\delta\boldsymbol{r}$ signifies a transformation under which the radius vectors $\boldsymbol{r}_a$ of all the particles (a being the number of the particle) receive the same increment $\delta\boldsymbol{r} : \boldsymbol{r}_a \to \boldsymbol{r}_a + \delta\boldsymbol{r}$. An arbitrary function $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$ of the coordinates of the particles, under such a transformation, becomes the function

$$\psi(\boldsymbol{r_1} + \delta\boldsymbol{r}, \boldsymbol{r_2} + \delta\boldsymbol{r}, \dots) = \psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots) + \delta\boldsymbol{r} \cdot \sum_a \nabla_a \psi$$

$$= (1 + \delta\boldsymbol{r} \cdot \nabla_a)\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$$

(∇_a denotes the operator of differentiation with respect to $\boldsymbol{r}_a$). The expression

$$1 + \delta\boldsymbol{r} \cdot \nabla_a$$

is the operator of an infinitely small displacement, which converts the function $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$ into the function

$$\psi(\boldsymbol{r}_1 + \delta\boldsymbol{r}, \boldsymbol{r}_2 + \delta\boldsymbol{r}, \dots).$$

The statement that some transformation does not change the Hamiltonian means that, if we make this transformation on the function $\widehat{H}\psi$, the result is the same as if we make it only on the function ψ and then apply the operator $\widehat{H}$. Mathematically, this can be written as follows. Let $\widehat{}$ be the operator which effects the transformation in question. Then we have $\widehat{O}(\widehat{H}\psi) = \widehat{H}(\widehat{O}\psi)$, whence

$$\widehat{O}\widehat{H} - \widehat{H}\widehat{O} = 0$$

i.e. the Hamiltonian must commute with the operator $\widehat{O}$.

In the case considered, the operator $\widehat{O}$ is the operator of an infinitely small displacement. Since the unit operator (the operator of multiplying by unity) commutes, of course, with any operator, and the constant factor $\delta\boldsymbol{r}$ can be taken in front of $\widehat{H}$, the condition $\widehat{O}\widehat{H} - \widehat{H}\widehat{O} = 0$ reduces here to

$$\left(\sum_a \nabla_a\right)\widehat{H} - \widehat{H}\left(\sum_a \nabla_a\right) = 0 \qquad (15.1)$$

As we know, the commutability of an operator (not containing the time explicitly) with $\widehat{H}$ means that the physical quantity corresponding to that operator is conserved. The quantity whose conservation for a closed system follows from the homogeneity of space is the momentum of the system (cf. Mechanics, §7). Thus the relation (15.1) expresses the law of conservation of momentum in quantum mechanics; the operator $\sum_a \nabla_a$ must correspond, apart from a constant factor, to the total momentum of the system, and each term ∇_a of the sum to the momentum of an individual particle.

The coefficient of proportionality between the operator $\widehat{\boldsymbol{p}}$ of the momentum of a particle and the operator ∇ can be determined by means of the passage to the limit of classical mechanics, and is $-\mathrm{i}\hbar$:

$$\widehat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla \tag{15.2}$$

or, in components,

$$\widehat{p}_x = -\mathrm{i}\hbar\frac{\partial}{\partial x}, \quad \widehat{p}_y = -\mathrm{i}\hbar\frac{\partial}{\partial y}, \quad \widehat{p}_z = -\mathrm{i}\hbar\frac{\partial}{\partial z}$$

Using the limiting expression (6.1) for the wave function, we have

$$\widehat{\boldsymbol{p}} = \mathrm{i}\hbar\frac{\mathrm{i}}{\hbar}\Psi\nabla\boldsymbol{S} = \Psi\nabla\boldsymbol{S}$$

i.e. in the classical approximation the effect of the operator reduces to multiplication by $\nabla\boldsymbol{S}$. The gradient $\nabla\boldsymbol{S}$ of the action is the classical momentum p of the particle (see Mechanics, §43).

It is easy to see that the operator (15.2) is Hermitian, as it should be. For, with arbitrary functions $\psi(x)$ and $\varphi(x)$ which vanish at infinity, we have

$$\int \varphi\widehat{p}_x\psi\,\mathrm{d}x = -\mathrm{i}\hbar\int \varphi\frac{\partial\psi}{\partial x}\,\mathrm{d}x = \mathrm{i}\hbar\int \psi\frac{\partial\varphi}{\partial x}\,\mathrm{d}x = \int \psi\widehat{p}_x^*\varphi\,\mathrm{d}x,$$

and this is the condition that the operator should be Hermitian.

Since the result of differentiating functions with respect to two different variables is independent of the order of differentiation, it is clear that the operators of the three components of momentum commute with one another:

$$\widehat{p}_x\widehat{p}_y - \widehat{p}_y\widehat{p}_x = 0, \quad \widehat{p}_x\widehat{p}_z - \widehat{p}_z\widehat{p}_x = 0 \quad \widehat{p}_y\widehat{p}_z - \widehat{p}_z\widehat{p}_y = 0 \tag{15.3}$$

This means that all three components of the momentum of a particle can simultaneously have definite values.

Let us find the eigenfunctions and eigenvalues of the momentum operators. They are determined by the vector equation

$$\mathrm{i}\hbar\nabla\psi = \boldsymbol{p}\psi \tag{15.4}$$

The solutions are of the form

$$\psi = \mathrm{C} \cdot \mathrm{e}^{\mathrm{i}\boldsymbol{pr}/\hbar} \tag{15.5}$$

where C is a constant. If all three components of the momentum are given simultaneously, we see that this completely determines the wave function of the particle. In other words, the quantities p_x, p_y, p_z form one of the possible complete sets of physical quantities for a particle. Their eigenvalues form a continuous spectrum extending from $-\infty$ to $+\infty$.

According to the rule (5.4) for normalizing the eigenfunctions of a continuous spectrum, the integral $\int \psi_{\boldsymbol{p'}}^* \psi_{\boldsymbol{p}} \mathrm{d}V$ taken over all space ($\mathrm{d}V = \mathrm{d}x\mathrm{d}y\mathrm{d}z$) must be equal to the delta function $\delta(\boldsymbol{p'} - \boldsymbol{p})$.[13] However, for reasons that will become clear from subsequent applications, it is more natural to normalize the eigenfunctions of the particle momentum by the delta function of the momentum difference divided by $2\pi\hbar$:

$$\int \psi_{\boldsymbol{p'}}^* \psi_{\boldsymbol{p}} \mathrm{d}V = \left(\frac{\boldsymbol{p'} - \boldsymbol{p}}{2\pi\hbar} \right)$$

or, equivalently,

$$\int \psi_{\boldsymbol{p'}}^* \psi_{\boldsymbol{p}} \mathrm{d}V = (2\pi\hbar)^3 \delta(\boldsymbol{p'} - \boldsymbol{p}) \tag{15.6}$$

(since each of the three factors in the three-dimensional delta function is $\delta[(p_x' - p_x)/2\pi\hbar] = 2\pi\hbar\delta(p_x'p_x)$, and so on).

The integration is effected by means of the formula[14]

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\alpha\xi} \mathrm{d}\xi = \delta(\alpha) \tag{15.7}$$

This shows that the constant in (15.5) is equal to unity if the normalization is according to (15.6):[15]

$$\psi_{\boldsymbol{p}} = \mathrm{e}^{\mathrm{i}\boldsymbol{pr}/\hbar} \tag{15.8}$$

---

[13] The three-dimensional function $\delta(\boldsymbol{a})$ of a vector $\boldsymbol{a}$ is defined as a product of delta functions of the components of the vector $\boldsymbol{a}$: $\delta(\boldsymbol{a}) = \delta(a_x)\delta(a_y)\delta(a_z)$.

[14] The conventional meaning of this formula is that the function on the left-hand side has the property (5.8) of the delta function. Substituting $\delta(x - a)$ in the form (15.7), we obtain from (5.8) the well-known Fourier integral formula

$$f(a) = \iint_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{i}\xi(x-a)}\mathrm{d}x\frac{\mathrm{d}\xi}{2\pi}$$

[15] Note that with this normalization the probability density $|\Psi|^2 = 1$, i.e. the function is normalised to "one particle per unit volume". This agreement of normalizations is, of course, no accident; see the last footnote to §48.

The expansion of an arbitrary wave function $\psi(\boldsymbol{r})$ of a particle in terms of the eigenfunctions $\psi_{\boldsymbol{p}}$ of its momentum operator is simply the expansion as a Fourier integral:

$$\psi(\boldsymbol{r}) = \int a(\boldsymbol{p})\psi_{\boldsymbol{p}}(\boldsymbol{r})\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3} = \int a(\boldsymbol{p})\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3} \tag{15.9}$$

(where $\mathrm{d}^3 p = \mathrm{d}p_x \mathrm{d}p_y \mathrm{d}p_z$). The expansion coefficients $a(\boldsymbol{p})$ are, according to formula (5.3),

$$a(\boldsymbol{p}) = \int \psi(\boldsymbol{r})\psi_{\boldsymbol{p}}^*(\boldsymbol{r})\mathrm{d}V = \int \psi(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\mathrm{d}V \tag{15.10}$$

The function $a(\boldsymbol{p})$ can be regarded (see §5) as the wave function of the particle in the "momentum representation";

$$|a(\boldsymbol{p})|^2 \frac{\mathrm{d}^3 p}{(2\pi\hbar)^3}$$

is the probability that the momentum has a value in the interval $\mathrm{d}^3 p$.

Just as the operator $\widehat{\boldsymbol{p}}$ corresponds to the momentum, determining its eigenfunctions in the coordinate representation, we can introduce the operator $\widehat{\boldsymbol{r}}$ of the coordinates of the particle in the momentum representation. It must be defined so that the mean value of the coordinates can be represented in the form

$$\overline{\boldsymbol{r}} = \int a^*(\boldsymbol{p})\widehat{\boldsymbol{r}}a(\boldsymbol{p})\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3} \tag{15.11}$$

On the other hand, this mean value is determined from the wave function $\psi(\boldsymbol{r})$ by

$$\overline{\boldsymbol{r}} = \int \psi^*\boldsymbol{r}\psi\mathrm{d}V.$$

Substituting $\psi(\boldsymbol{r})$ in the form (15.9) we have (integrating by parts)

$$\boldsymbol{r}\psi(\boldsymbol{r}) = \int \boldsymbol{r}a(\boldsymbol{p})\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3} = \int \mathrm{i}\hbar\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\frac{\partial a(\boldsymbol{p})}{\partial \boldsymbol{p}}\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3}$$

Using this expression and (15.10), we find

$$\overline{\boldsymbol{r}} = \iint \psi^*(\boldsymbol{r})\mathrm{i}\hbar\frac{\partial a(\boldsymbol{p})}{\partial \boldsymbol{p}}\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\mathrm{d}V\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3} = \int \mathrm{i}\hbar a^*(\boldsymbol{p})\frac{\partial a(\boldsymbol{p})}{\partial \boldsymbol{p}}\frac{\mathrm{d}^3 p}{(2\pi\hbar)^3}$$

Comparing with (15.11), we see that the radius vector operator in the momentum representation is

$$\widehat{\boldsymbol{r}} = \mathrm{i}\hbar\frac{\partial}{\partial \boldsymbol{p}} \tag{15.12}$$

The momentum operator in this representation reduces simply to multiplication by $\boldsymbol{p}$.

Finally, we shall express in terms of $\widehat{\boldsymbol{p}}$ the operator of a parallel displacement in space over any finite (not only infinitesimal) distance $\boldsymbol{a}$. By the definition of this operator $(\widehat{T}_{\boldsymbol{a}})$ we must have

$$\widehat{T}_{\boldsymbol{a}}\psi(\boldsymbol{r}) = \psi(\boldsymbol{r} + \boldsymbol{a}).$$

Expanding the function $\psi(\boldsymbol{r} + \boldsymbol{a})$ in a Taylor series, we have

$$\psi(\boldsymbol{r} + \boldsymbol{a}) = \psi(\boldsymbol{r}) + \boldsymbol{a} \cdot \frac{\partial \psi(\boldsymbol{r})}{\partial \boldsymbol{r}} + \dots,$$

or, introducing the operator $\widehat{\boldsymbol{p}} = \mathrm{i}\hbar\nabla$,

$$\psi(\boldsymbol{r} + \boldsymbol{a}) = \left[1 + \frac{\mathrm{i}}{\hbar}\boldsymbol{a} \cdot \widehat{\boldsymbol{p}} + \frac{1}{2}\left(\frac{\mathrm{i}}{\hbar}\boldsymbol{a} \cdot \widehat{\boldsymbol{p}} \right)^2 + \dots \right] \psi(\boldsymbol{r}).$$

The expression in brackets is the operator

$$\widehat{T}_{\boldsymbol{a}} = \exp\left(\frac{\mathrm{i}}{\hbar}\boldsymbol{a} \cdot \widehat{\boldsymbol{p}} \right) \tag{15.13}$$

This is the required operator of the *finite displacement.*

§ 16. Uncertainty relations

Let us derive the rules for commutation between momentum and coordinate operators. Since the result of successively differentiating with respect to one of the variables x, y, z and multiplying by another of them does not depend on the order of these operations, we have

$$\widehat{p}_x y - y\widehat{p}_x = 0, \quad \widehat{p}_x z - z\widehat{p}_x = 0 \tag{16.1}$$

and similarly for $\widehat{p}_y, \widehat{p}_z$.

To derive the commutation rule for $\widehat{p}_x$ and x, we write

$$(\widehat{p}_x x - x\widehat{p}_x)\psi = -\mathrm{i}\hbar\frac{\partial}{\partial x}(x\psi) + \mathrm{i}\hbar x\frac{\partial \psi}{\partial x} = \mathrm{i}\hbar\psi$$

We see that the result of the action of the operator $\widehat{p}_x x - x\widehat{p}_x$ reduces to multiplication by $\mathrm{i}\hbar$; the same is true, of course, of the commutation of $\widehat{p}_y$ with y and $\widehat{p}_z$ with z. Thus we have[16]

$$\widehat{p}_x - x\widehat{p}_x = -\mathrm{i}\hbar, \qquad \widehat{p}_y - y\widehat{p} = -\mathrm{i}\hbar, \qquad \widehat{p}_x x - x\widehat{p}_x = -\mathrm{i}\hbar \tag{16.2}$$

---

[16] These relations, discovered in matrix form by Heisenberg in 1925, formed the genesis of quantum mechanics.

All the relations (16.1) and (16.2) can be written jointly in the form

$$\widehat{p}_i x_k - x_k \widehat{p}_i = -i\hbar\delta_{ik}, \qquad (i, k = x, y, z) \tag{16.3}$$

Before going on to examine the physical significance of these relations and their consequences, we shall set down two formulae which will be useful later. Let $f(\boldsymbol{r})$ be some function of the coordinates. Then

$$\widehat{\boldsymbol{p}}f(\boldsymbol{r}) - f(\boldsymbol{r})\widehat{\boldsymbol{p}} = -i\hbar\nabla f \tag{16.4}$$

For

$$(\widehat{\boldsymbol{p}}f - f\widehat{\boldsymbol{p}})\psi = -i\hbar[\nabla(f\psi) - f\nabla\psi] = -i\hbar\psi\nabla f$$

A similar relation holds for the commutator of $\boldsymbol{r}$ with a function of the momentum operator:

$$f(\widehat{\boldsymbol{p}})\boldsymbol{r} - \boldsymbol{r}f(\widehat{\boldsymbol{p}}) = -i\hbar\frac{\partial f}{\partial \boldsymbol{p}} \tag{16.5}$$

It can be derived in the same way as (16.4) if we calculate in the momentum representation, using the expression (15.12) for the coordinate operators.

The relations (16.1) and (16.2) show that the coordinate of a particle along one of the axes can have a definite value at the same time as the components of the momentum along the other two axes; the coordinate and momentum component along the same axis, however, cannot exist simultaneously. In particular, the particle cannot be at a definite point in space and at the same time have a definite momentum $\boldsymbol{p}$.

Let us suppose that the particle is in some finite region of space, whose dimensions along the three axes are (of the order of magnitude of) Δx, Δy, Δz. Also, let the mean value of the momentum of the particle be $\boldsymbol{p}_0$. Mathematically, this means that the wave function has the form $\psi = u(\boldsymbol{r})e^{i\boldsymbol{p}_0\boldsymbol{r}/\hbar}$, where $u(\boldsymbol{r})$ is a function which differs considerably from zero only in the region of space concerned. We expand the function ψ in terms of the eigenfunctions of the momentum operator (i.e. as a Fourier integral). The coefficients $a(\boldsymbol{p})$ in this expansion are determined by the integrals (15.10) of functions of the form $u(\boldsymbol{r})e^{i(\boldsymbol{p}_0-\boldsymbol{p})\boldsymbol{r}/\hbar}$. If this integral is to differ considerably from zero, the periods of the oscillatory factor $e^{i(\boldsymbol{p}_0-\boldsymbol{p})\boldsymbol{r}/\hbar}$ must not be small in comparison with the dimensions Δx, Δy, Δz of the region in which the function $u(\boldsymbol{r})$ is different from zero. This means that $a(\boldsymbol{p})$ will be considerably different from zero only for values of $\boldsymbol{p}$ such that $(p_{0x} - p_x)\Delta x/\hbar \lesssim 1$, etc. Since $|a(\boldsymbol{p})|^2$ determines the probability of the various values of the momentum, the ranges of values of p_x, p_y, p_z in which $a(\boldsymbol{p})$ differs from zero are just those

in which the components of the momentum of the particle may be found, in the state considered. Denoting these ranges by Δp_x, Δp_y, Δp_z, we thus have

$$\Delta p_x \Delta x \sim \hbar, \qquad \Delta p_y \Delta y \sim \hbar, \qquad \Delta p_z \Delta z \sim \hbar \qquad (16.6)$$

These relations, known as the *uncertainty relations*, were obtained by Heisenberg in 1927.

We see that, the greater the accuracy with which the coordinate of the particle is known (i.e. the less Δx), the greater the uncertainty Δp_x in the component of the momentum along the same axis, and *vice versa*. In particular, if the particle is at some completely definite point in space ($\Delta x = \Delta y = \Delta z = 0$), then $\Delta p_x = \Delta p_y = \Delta p_z = \infty$. This means that all values of the momentum are equally probable. Conversely, if the particle has a completely definite momentum $\boldsymbol{p}$, then all positions of it in space are equally probable (this is seen directly from the wave function (15.8), whose squared modulus is quite independent of the coordinates).

If the uncertainties of the coordinates and momenta are specified by the standard deviations

$$\delta x = \sqrt{\overline{(x - \overline{x})^2}}, \quad \delta p_x = \sqrt{\overline{(x - \overline{x})^2}},$$

we can specify exactly the least possible value of their product (H. Weyl). Let us consider the one-dimensional case of a wave packet with wave function $\psi(x)$ depending on only one coordinate, and assume for simplicity that the mean values of x and p_x in this state are zero. We consider the obvious inequality

$$\int_{-\infty}^{\infty} \left| \alpha x \psi + \frac{\mathrm{d}\psi}{\mathrm{d}x} \right|^2 \mathrm{d}x \geqslant 0$$

where α is an arbitrary real constant. On calculating this integral, noticing that

$$\int_{-\infty}^{\infty} x^2 |\psi|^2 \mathrm{d}x = (\delta x)^2,$$

$$\int \left(x \frac{\mathrm{d}\psi^*}{\mathrm{d}x} \psi + x \psi^* \frac{\mathrm{d}\psi}{\mathrm{d}x} \right) \mathrm{d}x = \int x \frac{\mathrm{d}|\psi|^2}{\mathrm{d}x} \mathrm{d}x = - \int |\psi|^2 \mathrm{d}x = -1,$$

$$\int \frac{\mathrm{d}\psi^*}{\mathrm{d}x} \frac{\mathrm{d}\psi}{\mathrm{d}x} \mathrm{d}x = - \int \psi^* \frac{\mathrm{d}^2 \psi}{\mathrm{d}x^2} \mathrm{d}x = \frac{1}{\hbar^2} \int \psi^* \widehat{p}_x^2 \psi \mathrm{d}x = \frac{1}{\hbar^2} (\delta p_x)^2,$$

we obtain

$$\alpha^2 (\delta x)^2 - \alpha + \frac{(\delta p_x)^2}{\hbar^2} \geqslant 0$$

If this quadratic (in α) trinomial is positive for all α, its discriminant must be negative, which gives the inequality

$$\delta x \, \delta p_x \geqslant 0 \tag{16.7}$$

The least possible value of the product is $\hbar/2$, and occurs for wave packets with wave functions of the form

$$\psi = \frac{1}{(2\pi)^{1/4}\sqrt{\delta x}} \exp\left(\frac{i}{\hbar}p_0 x - \frac{x^2}{4(\delta x)^2}\right) \tag{16.8}$$

where p_0 and δx are constants. The probabilities of the various values of the coordinates in such a state are

$$|\psi|^2 = \frac{1}{\sqrt{2\pi}\delta x} \exp\left(-\frac{x^2}{2(\delta x)^2}\right)$$

and thus have a Gaussian distribution about the origin (the mean value $\overline{x} = 0$) with standard deviation δx. The wave function in the momentum representation is

$$a(p_x) = \int_{-\infty}^{\infty} \psi(x) \exp\left(-i\frac{p_x x}{\hbar}\right) dx$$

Calculation of the integral gives

$$a(p_x) = \text{const} \cdot \exp\left(-\frac{(\delta x)^2(p_x - p_0)^2}{\hbar^2}\right).$$

The distribution of probabilities of values of the momentum, $|a(p_x)|^2$, is also Gaussian about the mean value $\overline{p_x} = p_0$, with standard deviation $\delta p_x = \hbar/2\delta x$, so that the product $\delta p_x \delta x$ is indeed $\hbar/2$.

Finally, we shall derive another useful relation. Let f and g be two physical quantities whose operators obey the commutation rule

$$\widehat{f}\widehat{g} - \widehat{g}\widehat{f} = -i\hbar\widehat{c} \tag{16.9}$$

where $\widehat{c}$ is the operator of some physical quantity c. On the right-hand side of the equation the factor $\hbar$ is introduced in accordance with the fact that in the classical limit (i.e. as $\hbar \to 0$) all operators of physical quantities reduce to multiplication by these quantities and commute with one another. Thus, in the "quasi-classical" case, we can, to a first approximation, regard the right-hand side of equation (16.9) as being zero. In the next approximation,

the operator   can be replaced by the operator of simple multiplication by the quantity c. We then have

$$\widehat{f}\widehat{g} - \widehat{g}\widehat{f} = -\mathrm{i}\hbar c.$$

This equation is exactly analogous to the relation $\widehat{p}_x x - x\widehat{p}_x = -\mathrm{i}\hbar$, the only difference being that, instead of the constant $\hbar$, we have[17] the quantity $\hbar c$. We can therefore conclude, by analogy with the relation $\Delta x \Delta p_x \sim \hbar$, that in the quasi-classical case there is an uncertainty relation

$$\Delta f \Delta g \sim \hbar c \qquad (16.10)$$

for the quantities f and g.

In particular, if one of these quantities is the energy$(\widehat{f} \equiv \widehat{H})$ and the operator $(\widehat{g})$ of the other does not depend explicitly on the time, then by 9.2 $c = \dot{g}$, and the uncertainty relation in the quasi-classical case is

$$\Delta E \Delta g \sim \hbar \dot{g}. \qquad (16.11)$$

[17] The classical quantity c is the Poisson bracket of the quantities f and g; see the footnote in §9.

CHAPTER III

SCHRÖDINGER'S EQUATION

§ 17. Schrödinger's equation

The form of the wave equation of a physical sysytem is detemined by its Hamiltonian, which is therefore of fundamental significance in the whole mathematical fomulation of quantum mechanics.

The form of the Hamiltonian for a free particle is established by the general requirements imposed by the homogeneity and isotropy of space and by Galileo's relativity principle. In classical mechanics, these requirements lead to a quadratic dependence of the energy of the particle on its momentum: $E = p^2/2m$, where the constant m is called the mass of the particle (see Mechanics, §4). In quantum mechanics, the same requirements lead to a corresponding relation for the energy and momentum eigenvalues, these quantities being conserved and simultaneously measurable (for a free particle).

If the relation $E = p^2/2m$ holds for every eigenvalue of the energy and momentum, the same relation must hold for their operators also:

$$\widehat{H} = \frac{1}{2m}\left(\widehat{p}_x^2 + \widehat{p}_y^2 + \widehat{p}_z^2.\right) \tag{17.1}$$

Substituting here from (15.2), we obtain the Hamiltonian of a freely moving particle in the form

$$\widehat{H} = -\frac{\hbar}{2m}\Delta \tag{17.2}$$

where $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the Laplacian operator.

The Hamiltonian of a system of non-interacting particles is equal to the sum of the Hamiltonians of the separate particles:

$$\widehat{H} = -\frac{\hbar^2}{2}\sum_a \frac{\Delta_a}{m_a} \tag{17.3}$$

(the suffix a is the number of the particle; Δ_a is the Laplacian operator in which the differentiation is with respect to the coordinates of the ath particle).

In classical (non-relativistic) mechanics, the interaction of particles is described by an additive term in the Hamiltonian, the potential energy of the interaction $U(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$, which is a function of the coordinates of the particles. By adding a similar function to the Hamiltonian of the system, the interaction of particles can be represented in quantum mechanics[1] :

$$\widehat{H} = -\frac{\hbar^2}{2} \sum_a \frac{\Delta_a}{m_a} + U(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots). \tag{17.4}$$

The first term can be regarded as the operator of the kinetic energy and the second as that of the potential energy. In particular, the Hamiltonian for a single particle in an external field is

$$\widehat{H} = \frac{\widehat{p}^2}{2m} + U(x, y, z) = -\frac{\hbar^2}{2m}\Delta + U(x, y, z), \tag{17.5}$$

where $U(x, y, z)$ is the potential energy of the particle in the external field.

Substituting the expressions (17.2) to (17.5) in the general equation (8.1), we obtain the wave equations for the corresponding systems. We shall write out here the wave equation for a particle in an external field:

$$i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\Psi + U(x, y, z) \tag{17.6}$$

The equation (10.2), which determines the stationary states, takes the form

$$\frac{\hbar^2}{2m}\Delta\psi + [E - U(x, y, z)]\,\psi = 0 \tag{17.7}$$

The equations (17.6) and (17.7) were obtained by Schrödinger in 1926 and are called *Schrödinger's equations*.

For a free particle, equation (17.7) has the form

$$\frac{\hbar^2}{2m}\Delta\psi + E\psi = 0 \tag{17.8}$$

This equation has solutions finite in all space for any positive value of the energy E. For states with definite directions of motion, these solutions are

---

[1] This statement is, of course, not a logical consequence of the basic principles of quantum mechanics, and is to be regarded as a deduction from experiment.

eigenfunctions of the momentum operator, with $E = p^2/2m$. The complete (time-dependent) wave functions of such stationary states are

$$\Psi = \text{const} \cdot \exp\left(\frac{\mathrm{i}}{\hbar}\left(Et - \boldsymbol{pr}\right)\right). \tag{17.9}$$

Each such function, a *plane wave*, describes a state in which the particle has a definite energy E and momentum $\boldsymbol{p}$. The angular frequency of this wave is $E/\hbar$ and its wave vector $\boldsymbol{k} = \boldsymbol{p}/\hbar$; the corresponding wavelength $2\pi\hbar/p$ is called the *de Broglie wavelength* of the particle.[2]

The energy spectrum of a freely moving particle is thus found to be continuous, extending from zero to $+\infty$. Each of these eigenvalues (except $E = 0$) is degenerate, and the degeneracy is infinite. For there corresponds to every value of E, different from zero, an infinite number of eigenfunctions (17.9), differing in the direction of the vector $\boldsymbol{p}$, which has a constant absolute magnitude.

Let us enquire how the passage to the limit of classical mechanics occurs in Schrödinger's equation, considering for simplicity only a single particle in an external field. Substituting in Schrödinger's equation (17.6) the limiting expression (6.1) for the wave function, $\Psi = ae^{\mathrm{i}S/\hbar}$, we obtain, on performing the differentiation,

$$a\frac{\partial S}{\partial t} - \mathrm{i}\hbar\frac{\partial a}{\partial t} + \frac{a}{2m}(\nabla S)^2 - \frac{\mathrm{i}\hbar}{2m}a\Delta S - \frac{\mathrm{i}\hbar}{m}\nabla S\nabla a - \frac{\hbar^2}{2m}\Delta a + Ua = 0.$$

In this equation there are purely real and purely imaginary terms (we recall that S and a are real); equating each separately to zero, we obtain two equations

$$\frac{\partial S}{\partial t} + \frac{1}{2m}(\nabla S)^2 + U - \frac{\hbar^2}{2ma}\Delta a = 0,$$

$$\frac{\partial a}{\partial t} + \frac{a}{2m}\Delta S + \frac{1}{m}\nabla S\nabla a = 0.$$

Neglecting the term containing $\hbar$ in the first of these equations, we obtain

$$\frac{\partial S}{\partial t} + \frac{1}{2m}(\nabla S)^2 + U = 0, \tag{17.10}$$

that is; the classical Hamilton—Jacobi equation for the action S of a particle, as it should be. We see, incidentally, that, as $\hbar \to 0$, classical mechanics is valid as far as quantities of the first (and not only the zero) order in $\hbar$ inclusive.

---

[2] The idea of a wave related to a particle was first introduced by L. de Broglie in 1924.

The second equation obtained above, on multiplication by $2a$, can be rewritten in the form

$$\frac{\partial a^2}{\partial t} + \text{div}\left(a^2 \frac{\nabla S}{m}\right) = 0 \qquad (17.11)$$

This equation has an obvious physical meaning: a^2 is the probability density for finding the particle at some point in space($|\Psi|^2 = a^2$); $\nabla S/m = \boldsymbol{p}/m$ is the classical velocity $\boldsymbol{v}$ of the particle. Hence equation (17.11) is simply the equation of continuity, which shows that the probability density "moves" according to the laws of classical mechanics with the classical velocity $\boldsymbol{v}$ at every point.

PROBLEM

Find the transformation law for the wave function in a Galilean transformation.

SOLUTION. Let us apply the transformation to the wave function for free motion of a particle (a plane wave). Since any function Ψ can be expanded in plane waves, this will also give the transformation law for any wave function.

The plane waves in the frames of reference K and K' (K' moving with velocity V relative to K) are

$$\Psi(\boldsymbol{r}, t) = \text{const} \cdot \exp\left[i(\boldsymbol{pr} - Et)/\hbar\right],$$

$$\Psi'(\boldsymbol{r}', t) = \text{const} \cdot \exp\left[i(\boldsymbol{p'r'} - E't)/\hbar\right],$$

where $\boldsymbol{r} = \boldsymbol{r'} + \boldsymbol{V}t$; the particle momenta and energies in the two frames are related by

$$\boldsymbol{p} = \boldsymbol{p'} + m\boldsymbol{V}, E = E' + \boldsymbol{V}\boldsymbol{p'} + mV^2/2$$

(see Mechanics, §8). Substitution of these expressions in Ψ gives

$$\Psi(\boldsymbol{r}, t) = \Psi'(\boldsymbol{r}', t) \exp\left[\frac{i}{\hbar}\left(m\boldsymbol{V}\boldsymbol{r'} + \frac{mV^2}{2}t\right)\right] =$$

$$= \Psi'(\boldsymbol{r} - \boldsymbol{V}t, t) \exp\left[\frac{i}{\hbar}\left(m\boldsymbol{V}\boldsymbol{r} - \frac{mV^2}{2}t\right)\right]. \quad (1)$$

This formula does not contain the parameters of the free motion of the particle, and gives the required general transformation law for the wave function of any state of the particle. For a system of particles, the exponent in (1) contains a summation over the particles.

§ 18. The fundamental properties of Schrödinger's equation

The conditions which must be satisfied by solutions of Schrödinger's equation are very general in character. First of all, the wave function must be

single-valued and continuous in all space. The requirement of continuity is maintained even in cases where the field $U(x, y, z)$ itself has a surface of discontinuity. At such a surface both the wave function and its derivatives must remain continuous. The continuity of the derivatives, however, does not hold if there is some surface beyond which the potential energy U becomes infinite. A particle cannot penetrate at all into a region of space where $U = \infty$, i.e. we must have $\psi = 0$ everywhere in this region. The continuity of ψ means that ψ vanishes at the boundary of this region; the derivatives of ψ, however, in general are discontinuous in this case.

If the field $U(x, y, z)$ nowhere becomes infinite, then the wave function also must be finite in all space. The same condition must hold in cases where U becomes infinite at some point but does so only as $1/r^s$ with $s < 2$ (see also §35).

Let $U_{\min}$ be the least value of the function $U(x, y, z)$. Since the Hamiltonian of a particle is the sum of two terms, the operators of the kinetic energy $\widehat{T}$ and of the potential energy, the mean value $\overline{E}$ of the energy in any state is equal to the sum $\overline{T} + \overline{U}$. But all the eigenvalues of the operator $\widehat{T}$ (which is the Hamiltonian of a free particle) are positive; hence the mean value $\overline{T} \geqslant 0$. Recalling also the obvious inequality $\overline{U} > U_{\min}$, we find that $\overline{E} > U_{\min}$. Since this inequality holds for any state, it is clear that it is valid for all the eigenvalues of the energy:

$$E_n > U_{\min} \tag{18.1}$$

Let us consider a particle moving in an external field which vanishes at infinity; we define the function $U(x, y, z)$, in the usual way, so that it vanishes at infinity. It is easy to see that the spectrum of negative eigenvalues of the energy will then be discrete, i.e. all states with $E < 0$ in a field which vanishes at infinity are bound states. For, in the stationary states of a continuous spectrum, which correspond to infinite motion, the particle reaches infinity (see §10); however, at sufficiently large distances the field may be neglected, the motion of the particle may be regarded as free, and the energy of a freely moving particle can only be positive.

The positive eigenvalues, on the other hand, form a continuous spectrum and correspond to an infinite motion; for $E > 0$, Schrödinger's equation in general has no solutions (in the field concerned) for which the integral $\int |\psi|^2 \mathrm{d}V$ converges.[3]

Attention must be drawn to the fact that, in quantum mechanics, a particle in a finite motion may be found in those regions of space where $E < U$; the

---

[3] However, it must be mentioned that, for some particular mathematical forms of the function U (x, y, z) (which have no physical significance), a discrete set of values may be absent from the otherwise continuous spectrum.

probability $|\psi|^2$ of finding the particle tends rapidly to zero as the distance into such a region increases, yet it differs from zero at all finite distances. Here there is a fundamental difference from classical mechanics, in which a particle cannot penetrate into a region where $U > E$. In classical mechanics the impossibility of penetrating into this region is related to the fact that, for $E < U$, the kinetic energy would be negative, that is, the velocity would be imaginary. In quantum mechanics, the eigenvalues of the kinetic energy are likewise positive; nevertheless, we do not reach a contradiction here, since, if by a process of measurement a particle is localized at some definite point of space, the state of the particle is changed, as a result of this process, in such a way that it ceases in general to have any definite kinetic energy.

If $U(x, y, z) > 0$ in all space(and $U \to 0$ at infinity), then, by the inequality (18.1), we have $E_n > 0$. Since, on the other hand, for E > 0 the spectrum must be continuous, we conclude that, in this case, the discrete spectrum is absent altogether, i.e. only an infinite motion of the particle is possible.

Let us suppose that, at some point (which we take as origin), U tends to $-\infty$ in the manner

$$U \approx -\alpha r^{-s}, \quad \alpha > 0. \tag{18.2}$$

We consider a wave function finite in some small region (of radius r_0) about the origin, and equal to zero outside this region. The uncertainty in the values of the coordinates of a particle in such a wave packet is of the order of r_0; hence the uncertainty in the value of the momentum is $\sim \hbar/r_0$. The mean value of the kinetic energy in this state is of the order of $\hbar^2/mr_0^2$, and the mean value of the potential energy is $\sim -\alpha/r_0^s$. Let us first suppose that s < 2. Then the sum

$$h^2/mr_0^2 - \alpha/r_0^s$$

takes arbitrarily larpe negative values for sufficiently small r_0. If, however, the mean energy can take such values, this always means that the energy has negative eigenvalues which are arbitrarily large in absolute value. The motion of the particle in a very small region of space near the origin corresponds to the energy levels with large $|E|$. The "normal" state corresponds to a particle at the origin itself, i.e. the particle "falls" to the point $r = 0$.

If, however, s < 2, the energy cannot take arbitrarily large negative values. The discrete spectrum begins at some finite negative value. In this case the particle does not fall to the centre. It should be mentioned that, in classical mechanics, the fall of a particle to the centre would be possible in principle in any attractive field (i.e. for any positive s). The case $s = 2$ will be specially considered in §35.

Next, let us investigate how the nature of the energy spectrum depends on the behaviour of the field at large distances. We suppose that, as $r \to \infty$, the

potential energy, which is negative, tends to zero according to the power law (18.2) (r is now large in this formula), and consider a wave packet "filling" a spherical shell of large radius r_0 and thickness $\Delta r \ll r_0$. Then the order of magnitude of the kinetic energy is again $\hbar^2/m(\Delta r)^2$, and of the potential energy, $-\alpha/r_0^s$. We increase r_0, at the same time increasing Δr, in such a way that Δr increases proportionally to r_0. If $s < 2$, then the sum becomes negative for sufficiently large r_0. Hence it follows that there are stationary states of negative energy, in which the particle may be found, with a fair probability, at large distances from the origin. This, however, means that there are levels of arbitrarily small negative energy (it must be recalled that the wave functions rapidly tend to zero in the region of space where $U > E$). Thus, in this case, the discrete spectrum contains an infinite number of levels, which become denser and denser towards the level $E = 0$.

If the field diminishes as $-1/r^s$ at infinity, with $s > 2$, then there are not levels of arbitrarily small negative energy. The discrete spectrum terminates at a level with a non-zero absolute value, so that the total number of levels is finite.

Schrödinger's equation for the wave functions ψ of stationary states is real, as are the conditions imposed on its solution. Hence its solutions can always be taken as real.[4] The eigenfunctions of non-degenerate values of the energy are automatically real, apart from the unimportant phase factor. For ψ^* satisfies the same equation as ψ, and therefore must also be an eigenfunction for the same value of the energy; hence, if this value is not degenerate, ψ and ψ^* must be essentially the same, i.e. they can differ only by a constant factor (of modulus unity). The wave functions corresponding to the same degenerate energy level need not be real, however, but by a suitable choice of linear combinations of them we can always obtain a set of real functions.

The complete (time-dependent) wave functions Ψ are determined by an equation in whose coefficients i appears. This equation, however, retains the same form if we replace t in it by $-t$ and at the same time take the complex conjugate.[5] Hence we can always choose the functions Ψ in such a way that Ψ and Ψ^* differ only by the sign of the time.

As is well known, the equations of classical mechanics are unchanged by time reversal, i.e. when the sign of the time is reversed. In quantum mechanics, the symmetry with respect to the two directions of time is expressed, as we see, in the invariance of the wave equation when the sign of t is changed and Ψ is simultaneously replaced by Ψ^*. However, it must be recalled that

---

[4] These assertions are not valid for systems in a magnetic field.

[5] It is assumed that the potential energy U does not depend explicitly on the time: the system is either closed or in a constant (non-magnetic) field.

this symmetry here relates only to the equation, and not to the concept of measurement itself, which plays a fundamental part in quantum mechanics (as we have explained in detail in §7).

§ 19. The current density

In classical mechanics the velocity $\boldsymbol{v}$ of a particle is related to its momentum by $\boldsymbol{p} = m\boldsymbol{v}$. A similar relation holds between the corresponding operators in quantum mechanics, as we should expect. This is easily shown by calculating the operator by the general rule (9.2) for the differentiation of operators with respect to time:

$$\widehat{\boldsymbol{v}} = \frac{\mathrm{i}}{\hbar} \left(\widehat{H}\boldsymbol{r} - \boldsymbol{r}\widehat{H} \right).$$

Using the expression (17.5) for $\widehat{H}$ and formula (16.5), we obtain

$$\widehat{\boldsymbol{v}} = \boldsymbol{p}/m \tag{19.1}$$

Similar relations will clearly hold between the eigenvalues of the velocity and momentum, and between their mean values in any state.

The velocity, like the momentum of a particle, cannot have a definite value simultaneously with the coordinates. But the velocity multiplied by an infinitely short time interval $\mathrm{d}t$ gives the displacement of the particle in the time $\mathrm{d}t$. Hence the fact that the velocity cannot exist at the same time as the coordinates means that, if the particle is at a definite point in space at some instant, it has no definite position at an infinitely close subsequent instant.

We may notice a useful formula for the operator $\widehat{\dot{f}}$ of the derivative, with respect to time, of some quantity $f(r)$ which is a function of the radius vector of the particle. Bearing in mind that f commutes with $U(r)$, we find

$$\widehat{\dot{f}} = \frac{\mathrm{i}}{\hbar} \left(\widehat{H}f - f\widehat{H} \right) = \frac{\mathrm{i}}{2m\hbar} \left(\widehat{\boldsymbol{p}}^2 f - f\widehat{\boldsymbol{p}}^2 \right).$$

Using (16.4), we can write

$$\widehat{\boldsymbol{p}}^2 f = \widehat{\boldsymbol{p}}(f\widehat{\boldsymbol{p}} - \mathrm{i}\hbar\nabla f), \quad f\widehat{\boldsymbol{p}}^2 = (\widehat{\boldsymbol{p}}f - \mathrm{i}\hbar\nabla f)\widehat{\boldsymbol{p}}$$

Thus we obtain the required expression:

$$\widehat{\dot{f}} = \frac{1}{2m} \left(\widehat{\boldsymbol{p}}\nabla f + \nabla f \cdot \widehat{\boldsymbol{p}} \right) \tag{19.2}$$

Next, let us find the acceleration operator. We have

$$\widehat{\boldsymbol{v}} = \frac{i}{\hbar}\left(\widehat{H}\widehat{\boldsymbol{v}} - \widehat{\boldsymbol{v}}\widehat{H}\right) = \frac{i}{m\hbar}\left(\widehat{H}\widehat{\boldsymbol{p}} - \widehat{\boldsymbol{p}}\widehat{H}\right) = \frac{i}{m\hbar}\left(U\widehat{\boldsymbol{p}} - \widehat{\boldsymbol{p}}U\right)$$

Using formula (16.4), we find

$$m\widehat{\dot{\boldsymbol{v}}} = -\nabla U. \tag{19.3}$$

This operator equation is exactly the same in form as the equation of motion (Newton's equation) in classical mechanics.

The integral $\int|\Psi|^2 dV$, taken over some finite volume V, is the probability of finding the particle in this volume. Let us calculate the derivative of this probability with respect to time. We have

$$\frac{d}{dt}\int|\Psi|^2 dV = \int\left(\Psi\frac{\partial\Psi^*}{\partial t} + \Psi^*\frac{\partial\Psi}{\partial t}\right)dV = \frac{i}{\hbar}\int\left(\Psi\widehat{H}^*\Psi^* - \Psi^*\widehat{H}\Psi\right)dV.$$

Substituting here

$$\widehat{H} = \widehat{H}^* = -\frac{\hbar^2}{2m}\Delta + U(x,y,z)$$

and using the identity

$$\Psi\Delta\Psi^* - \Psi^*\Delta\Psi = \text{div}\left(\Psi\nabla\Psi^* - \Psi^*\nabla\Psi\right)$$

we obtain

$$\frac{d}{dt}\int|\Psi|^2 dV = -\int\text{div}\,\boldsymbol{j}\,dV$$

where $\boldsymbol{j}$ denotes the vector[6])

$$\boldsymbol{j} = \frac{i\hbar}{2m}(\Psi\nabla\Psi^* - \Psi^*\nabla\Psi) = \frac{1}{2m}(\Psi\widehat{\boldsymbol{p}}^*\Psi^* + \Psi^*\widehat{\boldsymbol{p}}\Psi). \tag{19.4}$$

The integral of $\text{div}\,\boldsymbol{j}$ can be transformed by Gauss's theorem into an integral over the closed surface which bounds the volume V:

$$\frac{d}{dt}\int|\Psi|^2 dV = -\oint\boldsymbol{j}\,d\boldsymbol{f} \tag{19.5}$$

---

[6]) If ψ is written as $|\psi|e^{i\alpha}$, then

$$\boldsymbol{j} = \frac{\hbar}{m}|\psi|^2\nabla\alpha \tag{19.4a}$$

It is seen from this that the vector $\boldsymbol{j}$ may be called the *probability current density* vector, or simply the *current density*. The integral of this vector over a surface is the probability that the particle will cross the surface during unit time. The vector $\boldsymbol{j}$ and the probability density $|\Psi|^2$ satisfy the equation

$$\frac{\partial |\Psi|^2}{\partial t} + \operatorname{div}\boldsymbol{j} = 0 \tag{19.6}$$

which is analogous to the classical equation of continuity.

The wave function of free motion (the plane wave (17.9)) can be normalized so as to describe a flow of particles with unit current density (in which, on average, one particle crosses a unit cross-section of the flow per unit time). This function is then

$$\Psi = \frac{1}{\sqrt{v}} \exp\left[-\frac{\mathrm{i}}{\hbar}(Et - \boldsymbol{pr})\right], \tag{19.7}$$

where v is the velocity of the particle, since substitution of this in (19.4) gives $\boldsymbol{j} = \boldsymbol{p}/mv$, i.e. a unit vector in the direction of the motion.

It is useful to show how the orthogonality of the wave functions of states with different energies follows immediately from Schrödinger's equation. Let ψ_m and ψ_n be two such functions; they satisfy the equations

$$-\frac{\hbar^2}{2m}\Delta\psi_m + U\psi_m = E_m\psi_m,$$

$$-\frac{\hbar^2}{2m}\Delta\psi_n^* + U\psi_n^* = E_n\psi_n^*.$$

We multiply the first of these by ψ_n^* and the second by ψ_m and subtract corresponding terms; this gives

$$(E_m - E_n)\psi_m\psi_n^* = \frac{\hbar^2}{2m}(\psi_m\Delta\psi_n^* - \psi_n^*\Delta\psi_m) = \frac{\hbar^2}{2m}\operatorname{div}(\psi_m\nabla\psi_n^* - \psi_n^*\nabla\psi_m).$$

If we now integrate both sides of this equation over all space, the right-hand side, on transformation by Gauss's theorem, reduces to zero, and we obtain

$$(E_m - E_n)\int \psi_m\psi_n^*\mathrm{d}V = 0,$$

whence, by the hypothesis $E_m \neq E_n$, there follows the required orthogonality relation

$$\int \psi_m\psi_n^*\mathrm{d}V = 0.$$

§ 20. The variational principle

SchrÖdinger's equation, in the general form $\widehat{H}\psi = E\psi$, can be obtained from, the variational principle

$$\delta \int \psi^*(\widehat{H} - E)\psi \mathrm{d}q = 0 \qquad (20.1)$$

Since ψ is complex, we can vary ψ and ψ^* independently. Varying ψ^*, we have

$$\int \delta\psi^*(\widehat{H} - E)\psi \mathrm{d}q = 0$$

whence, because $\delta\psi^*$ is arbitrary, we obtain the required equation $\widehat{H}\psi = E\psi$. The variation of ψ gives nothing different. For, varying ψ and using the fact that the operator $\widehat{H}$ is Hermitian, we have

$$\int \psi^*(\widehat{H} - E)\delta\psi \mathrm{d}q = \int \delta\psi(\widehat{H}^* - E)\psi^* \mathrm{d}q = 0$$

from which we obtain the complex conjugate equation $\widehat{H}^*\psi^* = E\psi^*$.

The variational principle (20.1) requires an unconditional extremum of the integral. It can be stated in a different form by regarding E as a Lagrangian multiplier in a problem with the conditional extremum requirement

$$\delta \int \psi^*\widehat{H}\psi \mathrm{d}q = 0 \qquad (20.2)$$

the condition being

$$\int \psi\psi^* \mathrm{d}q = 1 \qquad (20.3)$$

The least value of the integral in (20.2) (with the condition (20.3)) is the first eigenvalue of the energy, i.e. the energy E_0 of the normal state. The function ψ which gives this minimum is accordingly the wave function ψ_0 of the normal state.[7] The wave functions $\psi_n(n > 0)$ of the other stationary states correspond only to an extremum, and not to a true minimum of the integral.

In order to obtain, from the condition that the integral in (20.2) is a minimum, the wave function ψ_1 and the energy E_1 of the state next to the normal one, we must restrict our choice to those functions ψ which satisfy not only the normalization condition (20.3) but also the condition of orthogonality

---

[7] In the rest of this section we shall suppose the wave functions ψ to be real; they can always be so chosen (if there is no magnetic field).

with the wave function ψ_0 of the normal state: $\int \psi\psi_0 \mathrm{d}q = 0$. In general, if the wave functions $\psi_0, \psi_1, \ldots, \psi_{n-1}$ of the first n states (arranged in order of increasing energy) are known, the wave function of the next state gives a minimum of the integral in (20.2) with the additional conditions

$$\int \psi^2 \mathrm{d}q = 1, \quad \int \psi\psi_m \mathrm{d}q = 0 \quad m = 0, 1, 2, \ldots, n-1. \tag{20.4}$$

We shall give here some general theorems which can be proved from the variational principle.[8]

The wave function ψ_0 of the normal state does not become zero (or, as we say, has no *nodes*) for any finite values of the coordinates.[9] In other words, it has the same sign in all space. Hence, it follows that the wave functions $\psi_n (n > 0)$ of the other stationary states, being orthogonal to ψ_0, must have nodes (if ψ_n is also of constant sign, the integral $\int \psi_0\psi_n \mathrm{d}q$ cannot vanish).

Next, from the fact that ψ_0 has no nodes, it follows that the normal energy level cannot be degenerate. For, suppose the contrary to be true, and let ψ_0, ψ_0' be two different eigenfunctions corresponding to the level E_0. Any linear combination $c\psi_0 + c'\psi_0'$ will also be an eigenfunction; but by choosing the appropriate constants c, c', we can always make this function vanish at any given point in space, i.e. we can obtain an eigenfunction with nodes.

If the motion takes place in a bounded region of space, we must have $\psi = 0$ at the boundary of this region (see §18). To determine the energy levels, it is necessary to find, from the variational principle, the minimum value of the integral in (20.2) with this boundary condition. The theorem that the wave function of the normal state has no nodes means in this case that ψ_0 does not vanish anywhere inside this region.

We notice that, as the dimensions of the region containing the motion increase, all the energy levels E_n decrease; this follows immediately from the fact that an extension of the region increases the range of functions which can make the integral a minimum, and consequently the least value of the integral can only diminish.

The expression

$$\int \psi\widehat{H}\psi \mathrm{d}q = \int \left[-\sum_a \frac{\hbar^2}{2m_a} \psi\Delta_a\psi + U\psi^2 \right] \mathrm{d}q$$

[8] The proof of theorems concerning the zeros of eigenfunctions (see also §21) is given by M. A. Lavrent'ev and L. A. Lyusternik, *The Calculus of Variations (Kurs variatsionnogo ischislemya)*, 2nd edition, chapter IX, Moscow, 1950; R. Courant and D. Hilbert, *Methods of Mathematical Physics*, volume I, chapter VI, Interscience, New York, 1953.

[9] This theorem and its consequences are not in general valid for the wave functions of systems consisting of several identical particles (see the end of §63).

for the states of the discrete spectrum of a particle system may be transformed into another expression which is more convenient in practice. In the first term of the integrand we write

$$\psi \Delta_a \psi = \mathrm{div}_a(\psi \nabla_a \psi) - (\nabla_a \psi)^2.$$

The integral of $\mathrm{div}_a(\psi \nabla_a \psi)$ over all space is transformed into an integral over an infinitely distant closed surface, and since the wave functions of the states of a discrete spectrum tend to zero sufficiently rapidly at infinity, this integral vanishes. Thus

$$\int \psi \widehat{H} \psi \mathrm{d}q = \int \left[\sum_a \frac{\hbar^2}{2m_a} (\nabla_a \psi)^2 + U\psi^2 \right] \mathrm{d}q. \tag{20.5}$$

§ 21. General properties of motion in one dimension

If the potential energy of a particle depends on only one coordinate (x), then the wave function can be sought as the product of a function of y and z and a function of x only. The former of these is determined by Schrödinger's equation for free motion, and the second by the one-dimensional Schrödinger's equation

$$\frac{\mathrm{d}^2 \psi}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}[E - U(x)]\psi = 0 \tag{21.1}$$

Similar one-dimensional equations are evidently obtained for the problem of motion in a field whose potential energy is $U(x, y, z) = U_1(x) + U_2(y) + U_3(z)$, i.e. can be written as a sum of functions each of which depends on only one of the coordinates. In §§22–24 we shall discuss a number of actual examples of such "one-dimensional" motion. Here we shall obtain some general properties of the motion.

We shall show first of all that, in a one-dimensional problem, none of the energy levels of a discrete spectrum is degenerate. To prove this, suppose the contrary to be true, and let ψ_1 and ψ_2 be two different eigenfunctions corresponding to the same value of the energy. Since both of these satisfy the same equation (21.1), we have

$$\frac{\psi_1''}{\psi_1} = \frac{2m}{\hbar^2}(U - E) = \frac{\psi_2''}{\psi_2}$$

or $\psi_1''\psi_2 - \psi_1\psi_2'' = 0$ (the prime denotes differentiation with respect to x). Integrating this relation, we find

$$\psi_1'\psi_2 - \psi_1\psi_2' = \mathrm{const.} \tag{21.2}$$

Since $\psi_1 = \psi_2 = 0$ at infinity, the constant must be zero, and so

$$\psi_1'\psi_2 - \psi_1\psi_2' = 0,$$

or $\psi_1'/\psi_1 = \psi_2'/\psi_2$. Integrating again, we obtain $\psi_1 = \text{const} \cdot \psi_2$, i.e. the two functions are essentially identical.

The following theorem (called the *oscillation theorem*) may be stated for the wave functions $\psi_n(x)$ of a discrete spectrum. The function $\psi_n(x)$ corresponding to the $(n+1)$th eigenvalue E_n (the eigenvalues being arranged in order of magnitude), vanishes n times (for finite[10] values of x).

We shall suppose that the function $U(x)$ tends to finite limiting values as $x \to \pm\infty$ (though it need not be a monotonic function). We take the limiting value $U(+\infty)$ as the zero of energy (i.e. we put $U(+\infty) = 0$), and we denote $U(-\infty)$ by U_0, supposing that $U_0 > 0$. The discrete spectrum lies in the range of energy values for which the particle cannot move off to infinity; for this to be so, the energy must be less than both limiting values $U(\pm\infty)$, i.e. it must be negative:

$$E < 0, \tag{21.3}$$

and we must, of course, have in any case $E > U_{\min}$, i.e. the function $U(x)$ must have at least one minimum with $U_{\min} < 0$.

Let us now consider the range of positive energy values less than U_0:

$$0 < E < U_0 \tag{21.4}$$

In this range the spectrum will be continuous, and the motion of the particle in the corresponding stationary states will be infinite, the particle moving off towards $x = +\infty$. It is easy to see that none of the eigenvalues of the energy in this part of the spectrum is degenerate either. To show this, it is sufficient to notice that the proof given above (for the discrete spectrum) still holds if the functions ψ_1, ψ_2 are zero at only one infinity (in the present case they tend to zero as $x \to -\infty$).

For sufficiently large positive values of x, we can neglect $U(x)$ in Schrödinger's equation (21.1):

$$\psi'' + \frac{2m}{\hbar^2}E\psi = 0$$

This equation has real solutions in the form of a stationary plane wave

$$\psi = a\cos(kx + \delta) \tag{21.5}$$

---

[10]) If the particle can be found only on a limited segment of the x-axis, we must consider the zeros of $\psi_n(x)$ within that segment.

where a and δ are constants, and the *wave number* $k = p/\hbar = \sqrt{2mE}/\hbar$. This formula determines the asymptotic form (for $x \to +\infty$) of the wave functions of the non-degenerate energy levels in the range (21.4) of the continuous spectrum. For large negative values of x, Schrödinger's equation is

$$\psi'' - \frac{2m}{\hbar^2}(U_0 - E)\psi = 0$$

The solution which does not become infinite as $x \to -\infty$ is

$$\psi = b\mathrm{e}^{\varkappa x}, \quad \varkappa = \frac{1}{2m}\sqrt{U_0 - E}. \tag{21.6}$$

This is the asymptotic form of the wave function as $x \to \infty$. Thus the wave function decreases exponentially in the region where $E < U$.

Finally, for

$$E > U_0 \tag{21.7}$$

the spectrum will be continuous, and the motion will be infinite in both directions. In this part of the spectrum all the levels are doubly degenerate. This follows from the fact that the corresponding wave functions are determined by the second-order equation (21.1), and both of the two independent solutions of this equation satisfy the necessary conditions at infinity (whereas, for instance, in the previous case one of the solutions became infinite as $x \to -\infty$, and therefore had to be rejected). The asymptotic form of the wave function as $x \to +\infty$ is

$$\psi = a_1 \mathrm{e}^{\mathrm{i}kx} + a_2 \mathrm{e}^{-\mathrm{i}kx} \tag{21.8}$$

and similarly for $x \to -\infty$. The term $\mathrm{e}^{\mathrm{i}kx}$ corresponds to a particle moving to the right, and $\mathrm{e}^{-\mathrm{i}kx}$ corresponds to one moving to the left.

Let us suppose that the function $U(x)$ is even $[U(x) = U(x)]$. Then Schrödinger's equation (21.1) is unchanged when the sign of the coordinate is reversed. It follows that, if $\psi(x)$ is some solution of this equation, then $\psi(-x)$ is also a solution, and coincides with $\psi(x)$ apart from a constant factor: $\psi(-x) = c\psi(x)$. Changing the sign of x again, we obtain $\psi(x) = c^2\psi(x)$, whence $c = \pm 1$. Thus, for a potential energy which is symmetrical (relative to $x = 0$), the wave functions of the stationary states must be either even $[\psi(-x) = \psi(x)]$ or odd $[\psi(-x) = -\psi(x)]$.[11] In particular, the wave function

---

[11] In this discussion it is assumed that the stationary state is not degenerate, i.e. the motion is not infinite in both directions. Otherwise, when the sign of x is changed, two wave functions belonging to the energy level concerned may be transformed into each other. In this case, however, although the wave functions of the stationary states need not be even or odd, they can always be made so (by choosing appropriate linear combinations of the original functions).

of the ground state is even, since it cannot have a node, while an odd function always vanishes for $x = 0$ $[\psi(0) = -\psi(0) = 0]$.

To normalize the wave functions of one-dimensional motion (in a continuous spectrum), there is a simple method of determining the normalization coefficient directly from the asymptotic expression for the wave function for large values of $|x|$.

Let us consider the wave function of a motion infinite in one direction, $x \to +\infty$. The normalization integral diverges as $x \to \infty$ (as $x \to -\infty$, the function decreases exponentially, so that the integral rapidly converges). Hence, to determine the normalization constant, we can replace ψ by its asymptotic value (for large positive x), and perform the integration, taking as the lower limit any finite value of x, say zero; this amounts to neglecting a finite quantity in comparison with an infinite one. We shall show that the wave function normalized by the condition

$$\int \psi_p^* \psi_{p'} \mathrm{d}x = \delta\left(\frac{p - p'}{2\pi\hbar}\right) = 2\pi\hbar\delta(p - p'), \qquad (21.9)$$

where p is the momentum of the particle at infinity, must have the asymptotic form (21.5) with $a = 2$:

$$\psi_p \approx 2\cos(kx + \delta) = \mathrm{e}^{\mathrm{i}(kx+\delta)} + \mathrm{e}^{-\mathrm{i}(kx+\delta)} \qquad (21.10)$$

Since we do not intend to verify the orthogonality of the functions corresponding to different p, on substituting the functions (21.10) in the normalization integral we shall suppose the momenta p and p' to be arbitrarily close; we can therefore put $\delta = \delta'$ (in general δ is a function of p). Next, we retain in the integrand only those terms which diverge for $p = p'$; in other words, we omit terms containing the factor $\mathrm{e}^{\pm\mathrm{i}(k+k')x}$. Thus we obtain

$$\int \psi_p^* \psi_{p'} \mathrm{d}x = \int_0^{+\infty} \mathrm{e}^{\mathrm{i}(k'-k)x}\mathrm{d}x + \int_0^{+\infty} \mathrm{e}^{-\mathrm{i}(k'-k)x}\mathrm{d}x = \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}(k'-k)x}\mathrm{d}x$$

which, from (15.7), is the same as (21.9).

The change to normalization by the delta function of energy is effected, in accordance with (5.14), by multiplying ψ_p by

$$\left(\frac{\mathrm{d}(p/2\pi\hbar)}{\mathrm{d}E}\right)^{1/2} = \frac{1}{\sqrt{2\pi\hbar v}}$$

where v is the velocity of the particle at infinity. Thus

$$\psi_E = \frac{1}{\sqrt{2\pi\hbar v}}\psi_p = \frac{1}{\sqrt{2\pi\hbar v}}(\mathrm{e}^{\mathrm{i}(kx+\delta)} + \mathrm{e}^{-\mathrm{i}(kx+\delta)}) \qquad (21.11)$$

The current density is $1/(2\pi\hbar)$ in each of the travelling waves that make up the stationary wave (21.11). Thus we can formulate the following rule for the normalization of the wave function for a motion infinite in one direction by the delta function of energy: having represented the asymptotic expression for the wave function in the form of a sum of two plane waves travelling in opposite directions, we must choose the normalization coefficient in such a way that the current density in the wave travelling towards (or away from) the origin is $1/(2\pi\hbar)$.

Similarly, we can obtain an analogous rule for normalizing the wave functions of a motion infinite in both directions. The wave function will be normalized by the delta function of energy if the sum of the probability currents in the waves travelling towards the origin from $x = +\infty$ and $x = -\infty$ is $1/(2\pi\hbar)$.

§ 22. The potential well

As a simple example of one-dimensional motion, let us consider motion in a square potential well, i.e. in a field where $U(x)$ has the form shown in FIG.1: $U(x) = 0$ for $0 < x < a$, $U(x) = U_0$ for $x < 0$ and $x > a$. It is evident a priori that for $E < U_0$ the spectrum will be discrete, while for $E > U_0$ we have a continuous spectrum of doubly degenerate levels.

In the region $0 < x < a$ we have Schrödinger's equation

$$\psi'' + \frac{2m}{\hbar^2}E\psi = 0 \qquad (22.1)$$

(the prime denotes differentiation with respect to x), while in the region outside the well

FIG.1

$$\psi'' + \frac{2m}{\hbar^2}(E - U_0)\psi = 0 \qquad (22.2)$$

For $x = 0$ and $x = a$ the solutions of these equations must be continuous together with their derivatives, while for $x = \pm\infty$ the solution of equation (22.2) must remain finite (for the discrete spectrum when $E < U_0$, it must vanish).

For $E < U_0$, the solution of equation (22.2) which vanishes at infinity is $\psi = \text{const} \cdot e^{\mp\varkappa x}$, where

$$\varkappa = \frac{1}{\hbar}\sqrt{2m(U_0 - E)}; \qquad (22.3)$$
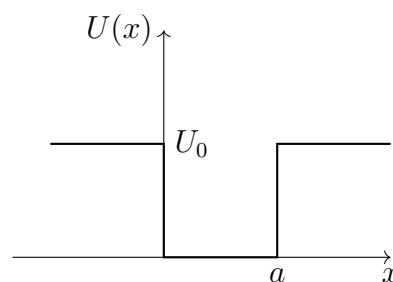
the signs $-$ and $+$ in the exponent refer to the regions $x > a$ and $x < 0$ respectively. The probability $|\psi|^2$ of finding the particle decreases exponentially in the region where $E < U(x)$. Instead of the continuity of ψ and ψ' at the edge of the potential well, it is convenient to require the continuity of ψ and of its logarithmic derivative ψ'/ψ. Taking account of (22.3), we obtain the boundary condition in the form

$$|\psi'|/\psi = \mp\varkappa \tag{22.4}$$

We shall not pause here to determine the energy levels in a well of arbitrary depth U_0 (see Problem 2), and shall analyse fully only the limiting case of infinitely high walls ($U_0 \to \infty$).

For $U_0 = \infty$, the motion takes place only between the points $x = 0$ and $x = a$ and, as was pointed out in §18, the boundary condition at these points is

$$\psi = 0 \tag{22.5}$$

(It is easy to see that this condition is also obtained from the general condition (22.4). For, when $U_0 \to \infty$, we have also $\varkappa \to \infty$ and hence $\psi'/\psi \to \infty$; since ψ' cannot become infinite, it follows that $\psi = 0$.) We seek a solution of equation (22.1) inside the well in the form

$$\psi = c\sin(kx + \delta), \quad k = \frac{\sqrt{2mE}}{\hbar}. \tag{22.6}$$

The condition $\psi = 0$ for $x = 0$ gives $\delta = 0$, and then the same condition for $x = a$ gives $\sin ka = 0$, whence $ka = n\pi$, n being a positive integer,[12] or

$$E_n = \frac{\pi^2\hbar^2}{2ma^2}n^2, \quad n = 1, 2, 3, \ldots \tag{22.7}$$

This determines the energy levels of a particle in a potential well. The normalized wave functions of the stationary states are

$$\psi_n = \sqrt{\frac{2}{n}}\sin\left(\frac{\pi n}{a}x\right). \tag{22.8}$$

From these results we can immediately write down the energy levels for a particle in a rectangular "potential box", i.e. for three-dimensional motion in a field whose potential energy $U = 0$ for $0 < x < a$, $0 < y < b$, $0 < z < c$ and $U = \infty$ outside this region. In fact, these levels are given by the sums

$$E_{n_1 n_2 n_3} = \frac{\pi^2\hbar^2}{2m}\left(\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2}\right), \quad n_1, n_2, n_3 = 1, 2, 3, \ldots, \tag{22.9}$$

---

[12] For $n = 0$ we should have $\psi = 0$ identically.

and the corresponding wave functions by the products

$$\psi_{n_1 n_2 n_3} = \sqrt{\frac{8}{abc}} \sin\left(\frac{\pi n_1}{a}x\right) \sin\left(\frac{\pi n_2}{b}y\right) \sin\left(\frac{\pi n_3}{c}z\right). \qquad (22.10)$$

It may be noted that the energy E_0 of the ground state is, by (22.7) or (22.9), of the order of $\hbar^2/ml^2$, where l is the linear dimension of the region in which the particle moves. This result is in accordance with the uncertainty relation; when the uncertainty in the coordinate is $\sim l$, the uncertainty in the momentum, and therefore the order of magnitude of the momentum itself, is $\sim \hbar/l$. The corresponding energy is $\sim (\hbar/l)^2/m$.

PROBLEMS

1. Determine the probability distribution for various values of the momentum for the normal state of a particle in an infinitely deep square potential well.

SOLUTION. The coefficients $a(p)$ in the expansion of the function ψ_1 (22.8) in terms of the eigenfunctions of the momentum are

$$a(p) = \int \psi_p^* \psi_1 \mathrm{d}x = \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{\pi}{a}x\right) \exp\left(-\mathrm{i}\frac{px}{\hbar}\right) \mathrm{d}x.$$

Calculating the integral and squaring its modulus, we obtain the required probability distribution:

$$|a(p)|^2 \frac{\mathrm{d}p}{2\pi\hbar} = \frac{4\pi\hbar^3 a}{(p^2 a^2 - \pi^2\hbar^2)^2} \cos^2\frac{pa}{2\hbar} \mathrm{d}p.$$

2. Determine the energy levels for the potential well shown in FIG.2.

SOLUTION. The spectrum of energy values $E < U_1$, which we shall consider, is discrete. In the region $x < 0$ the wave function is

FIG.2

$$\psi = c_1 \mathrm{e}^{\varkappa_1 x}, \quad \varkappa_1 = (1/\hbar)\sqrt{2m(U_1 - E)},$$

while in the region $x > a$

$$\psi = c_2 \mathrm{e}^{-\varkappa_2 x}, \quad \varkappa_2 = (1/\hbar)\sqrt{2m(U_2 - E)}.$$
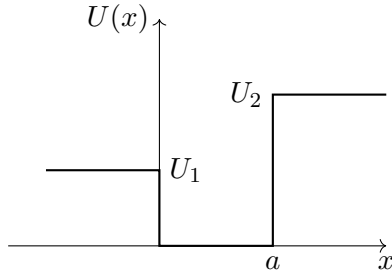
Inside the well ($0 < x < a$) we look for ψ in the form

$$\psi = c \sin(kx + \delta), \quad k = \sqrt{2mE}/\hbar.$$

The condition of the continuity of ψ'/ψ at the edges of the well gives the equations

$$k \cot\delta = \varkappa_1 = \sqrt{\frac{2m}{\hbar^2}U_1 - k^2}, \quad k \cot(ka + \delta) = -\varkappa_2 = -\sqrt{\frac{2m}{\hbar^2}U_2 - k^2},$$

or

$$\sin \delta = \frac{k\hbar}{\sqrt{2mU_1}}, \quad \sin(ka + \delta) = -\frac{k\hbar}{\sqrt{2mU_2}}.$$

Eliminating δ, we obtain the transcendental equation

$$ka = n\pi - \arcsin \frac{k\hbar}{\sqrt{2mU_1}} - \arcsin \frac{k\hbar}{\sqrt{2mU_2}} \tag{1}$$

(where $n = 1, 2, 3, \ldots$, and the values of the inverse sine are taken between 0 and $\pi/2$), whose roots determine the energy levels $E = k^2\hbar^2/2m$. For each n there is in general one root; the values of n number the levels in order of increasing energy.

Since the argument of the inverse sine cannot exceed unity, it is clear that the values of k can lie only in the range from 0 to $\sqrt{2mU_1}/\hbar$. The left-hand side of equation (1) increases monotonically with k, and the right-hand side decreases monotonically. Hence it is necessary, for a root of equation (1) to exist, that for the right-hand side should be less than the left-hand side. In particular, the inequality

$$a\frac{\sqrt{2mU_1}}{\hbar} \geqslant \frac{\pi}{2} - \arcsin \sqrt{\frac{U_1}{U_2}}, \tag{2}$$

which is obtained for $n = 1$, is the condition that at least one energy level exists in the well. We tee that for given and unequal $U_1 \neq U_2$, there are always widths a of the well which are so small that there is no discrete energy level. For $U_1 = U_2$, the condition (2) is evidently always satisfied.

For $U_1 = U_2 = U_0$ (a symmetrical well), equation (1) reduces to

$$\arcsin \frac{\hbar k}{\sqrt{2mU_0}} \frac{n\pi - ka}{2}. \tag{3}$$

Introducing the variable $\xi = ka/2$, we obtain for odd n the equation

$$\cos \xi = \pm\gamma\xi, \quad \gamma = \frac{\hbar}{a}\sqrt{\frac{2}{mU_0}}, \tag{4}$$

and those roots of this equation must be taken for which $\tan \xi > 0$. For even n we obtain the equation

$$\sin \xi = \pm\gamma\xi, \tag{5}$$

and we must take those roots for which $\tan \xi < 0$. The roots of these two equations determine the energy levels $E = 2\xi^2\hbar^2/ma^2$. The number of levels is finite when $\gamma \neq 0$.

In particular, for a shallow well in which $U_0 \ll \hbar^2/ma^2$, we have $\gamma \gg 1$ and equation (5) has no root. Equation (4) has one root (with the upper sign on the right-hand side), $\xi \approx (1/\gamma)(1 - 1/2\gamma^2)$. Thus the well contains only one energy level,

$$E_0 \approx U_0 - \frac{ma^2}{2\hbar^2}U_0^2$$

which is near the top of the well.

3. Determine the pressure exerted on the walls of a rectangular "potential box" by a particle inside it.

SOLUTION. The force on the the wall perpendicular to the x-axis is the mean value of the derivative $-\partial H/\partial a$ of the Hamilton's function of the particle with respect to the length of the box in the direction of the x-axis. The pressure is obtained by dividing this force by the area bc of the wall. According to the formula (11.16), the required mean value is found by differentiating the eigenvalue (22.9) of the energy. The result is

$$p^{(x)} = \frac{\pi^2 \hbar^2}{ma^3 bc} n_1^2.$$

§ 23. The linear oscillator

Let us consider a particle executing small oscillations in one dimension (what is called a *linear oscillator*). The potential energy of such a particle is $m\omega^2 x^2/2$, where ω is, in classical mechanics, the characteristic (angular) frequency of the oscillations. Accordingly, the Hamiltonian of the oscillator is

$$\widehat{H} = \frac{\widehat{p}^2}{2m} + \frac{m\omega^2 x^2}{2}. \tag{23.1}$$

Since the potential energy becomes infinite for $x = \pm\infty$, the particle can have only a finite motion, and the energy eigenvalue spectrum is entirely discrete.

Let us determine the energy levels of the oscillator, using the matrix method[13] . We shall start from the equations of motion in the form (19.3); in this case they give

$$\widehat{\ddot{x}} + \omega^2 x = 0. \tag{23.2}$$

In matrix form, this equation reads

$$(\ddot{x})_{mn} + \omega^2 x_{mn} = 0.$$

For the matrix elements of the acceleration we have, according to (11.8), $(\ddot{x})_{mn} = i\omega_{mn}(\dot{x})_{mn} = -\omega_{mn}^2 x_{mn}$. Hence we obtain

$$(\omega_{mn}^2 - \omega^2)x_{mn} = 0.$$

Hence it is evident that all the matrix elements x_{mn} vanish except those for which $\omega_{mn} = \omega$ or $\omega_{mn} = -\omega$. We number all the stationary states so that

---

[13] This was done by Heisenberg in 1925, before Schrödinger's discovery of the wave equation.

the frequencies $\pm\omega$ correspond to transitions $n \to n \mp 1$, i.e. $\omega_{n,n\mp1} = \pm\omega$. Then the only non-zero matrix elements are $x_{n,n\pm1}$.

We shall suppose that the wave functions ψ_n are taken real. Since x is a real quantity, all the matrix elements x_{mn} are real. The Hermitian condition (11.10) now shows that the matrix x_{mn} is symmetrical:

$$x_{mn} = x_{nm}.$$

To calculate the matrix elements of the coordinate which are different from zero, we use the commutation rule

$$\widehat{\dot{x}}\widehat{x} - \widehat{x}\widehat{\dot{x}} = -\mathrm{i}\frac{\hbar}{m},$$

written in the matrix form

$$(\dot{x}x)_{mn} - (x\dot{x})_{mn} = -\frac{\mathrm{i}\hbar}{m}\delta_{mn}$$

By the matrix multiplication rule (11.12) we hence have for $m = n$

$$\mathrm{i}\sum_l (\omega_{nl}x_{nl}x_{ln} - x_{nl}\omega_{ln}x_{ln}) = 2\mathrm{i}\sum_l \omega_{nl}x_{nl}^2 = -\mathrm{i}\frac{\hbar}{m}.$$

In this sum, only the terms with $l = n\pm1$ are different from zero, so that we have

$$(x_{n+1,n})^2 - (x_{n,n-1})^2 = \frac{\hbar}{2m\omega}. \tag{23.3}$$

From this equation we deduce that the quantities $(x_{n+1,n})^2$ form an arithmetic progression, which is unbounded above, but is certainly bounded below, since it can contain only positive terms. Since we have as yet fixed only the relative positions of the numbers n of the states, but not their absolute values, we can arbitrarily choose the value of n corresponding to the first (normal) state of the oscillator, and put this value equal to zero. Accordingly $x_{0,-1}$ must be regarded as being zero identically, and the application of equations (23.3) with $n = 0, 1, \cdots$ successively leads to the result

$$(x_{n,n-1})^2 = \frac{n\hbar}{2m\omega}.$$

Thus we finally obtain the following expression for the matrix elements of the coordinate which are different from zero:[14]

$$x_{n,n-1} = x_{n-1,n} = \sqrt{\frac{n\hbar}{2m\omega}} \tag{23.4}$$

---

[14] We choose the indeterminate phases α_n (see the second footnote to §11) so as to obtain the plus sign in front of the radical in all the matrix elements (23.4). Such a choice is always possible for a matrix in which only those elements are different from zero which correspond to transitions between states with adjacent numbers.

The matrix of the operator $\widehat{H}$ is diagonal, and the matrix elements H_{nn} are the required eigenvalues E_n of the energy of the oscillator. To calculate them, we write

$$H_{nn} = E_n = \frac{m}{2} \left[\left(\dot{x}^2 \right)_{nn} + \omega^2 (x^2)_{nn} \right] =$$

$$= \frac{m}{2} \left[\sum_l i\omega_{nl} x_{nl} i\omega_{ln} x_{ln} + \omega^2 \sum_l x_{nl} x_{ln} \right] = \frac{m}{2} \sum_l \left(\omega^2 + \omega_{nl}^2 \right) x_{ln}^2.$$

In the sum over l, only the terms with $l = n \pm 1$ are different from zero; substituting (23.4), we obtain

$$E_n = (n + 1/2)\hbar\omega, \quad n = 0, 1, 2, \ldots \tag{23.5}$$

Thus the energy levels of the oscillator lie at equal intervals of $\hbar\omega$ from one another. The energy of the normal state $(n = 0)$ is $\hbar\omega/2$; we call attention to the fact that it is not zero.

The result (23.5) can also be obtained by solving Schrödinger's equation. For an oscillator, this has the form

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{2m}{\hbar^2} \left(E - \frac{m\omega^2 x^2}{2} \right) \psi = 0 \tag{23.6}$$

Here it is convenient to introduce, instead of the coordinate x, the dimensionless variable ξ by the relation

$$\xi = \sqrt{\frac{m\omega}{\hbar}} x. \tag{23.7}$$

Then we have the equation

$$\psi'' + \left(\frac{2E}{\hbar\omega} - \xi^2 \right) \psi = 0; \tag{23.8}$$

here the prime denotes differentiation with respect to ξ.

For large ξ, we can neglect $2E/\hbar\omega$ in comparison with ξ^2; the equation $\psi'' = \xi^2\psi$ has the asymptotic integrals $\psi = \mathrm{e}^{\pm\xi^2/2}$ (for differentiation of this function gives $\psi'' = \xi^2\psi$ on neglecting terms of order less than that of the term retained). Since the wave function ψ must remain finite as $\xi \to \pm\infty$, the index must be taken with the minus sign. It is therefore natural to make in equation (23.8) the substitution

$$\psi = \mathrm{e}^{-\xi^2/2}\chi(\xi). \tag{23.9}$$

For the function we obtain the equation (with the notation $2E/1 = 2n$; since we already know that $E > 0$, we have $n > -1/2$)

$$\chi'' - 2\xi\chi' + 2n\chi = 0, \tag{23.10}$$

where the function must be finite for all finite ξ, and for $\xi \to \pm\infty$ must not tend to infinity more rapidly than every finite power of ξ (in order that the function ψ should tend to zero).

Such solutions of equation (23.10) exist only for positive integral (and zero) values of n (see §a of the Mathematical Appendices); this gives the eigenvalues (23.5) for the energy, which we know already. The solutions of equation (23.10) corresponding to various integral values of n are

$$\chi = \text{const} \cdot H_n(\xi),$$

where $H_n(\xi)$ are what are called *Hermite polynomials*; these are polynomials of the nth degree in ξ, defined by the formula

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n e^{-\xi^2}}{d\xi^n} \tag{23.11}$$

Determining the constants so that the functions ψ_n satisfy the normalization condition

$$\int_{-\infty}^{+\infty} \psi_n^2(x) dx = 1$$

we obtain (see (a.7))

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) H_n\left(x\sqrt{\frac{m\omega}{\hbar}}\right). \tag{23.12}$$

Thus the wave function of the normal state is

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right). \tag{23.13}$$

It has no zeros for finite x, which is as it should be.

By calculating the integrals $\int_{-\infty}^{+\infty} \psi_n\psi_m\xi d\xi$, we can determine the matrix elements of the coordinate; this calculation leads, of course, to the same values (23.4).

Finally, we shall show how the wave functions ψ_n may be calculated by the matrix method. We notice that, in the matrices of the operators $\widehat{\dot{x}} \pm i\omega\widehat{x}$, the only elements different from zero are

$$(\dot{x} - i\omega x)_{n,n-1} = -(\dot{x} + i\omega)_{n,n-1} = -i\sqrt{\frac{2\omega\hbar n}{m}} \tag{23.14}$$

Using the general formula (11.11), and taking into account the fact that $\psi_1 \equiv 0$, we conclude that

$$(\widehat{\dot{x}} - i\omega x)\psi_0 = 0$$

After substituting the expression $\widehat{\dot{x}} = -i\frac{\hbar}{m}\frac{d}{dx}$, we obtain the equation

$$\frac{d\psi_0}{dx} = -\frac{m\omega}{\hbar}x\psi_0$$

whose normalized solution is (23.13). And, since

$$(\widehat{\dot{x}} + i\omega\widehat{x})\psi_{n-1} = (\dot{x} + i\omega x)_{n,n-1}\psi_n = i\sqrt{\frac{2\omega\hbar n}{m}}\psi_n,$$

we obtain the recurrence formula

$$\psi_n = \sqrt{\frac{m}{2\omega\hbar n}}\left(-\frac{\hbar}{m}\frac{d}{dx} + \omega x\right)\psi_{n-1} = \frac{1}{\sqrt{2n}}\left(-\frac{d}{d\xi} + \xi\right)\psi_{n-1} =$$

$$= -\frac{1}{\sqrt{2n}}\exp\frac{\xi^2}{2}\frac{d}{d\xi}\left(\exp\left(-\frac{\xi^2}{2}\right)\psi_{n-1}\right),$$

when this is applied n times to the function (23.13), we obtain the expression (23.12) for the normalized functions ψ_n.

PROBLEMS

1. Determine the probability distribution of the various values of the momentum for an oscillator.

SOLUTION. Instead of expanding the wave function of the stationary state in terms of the eigenfunctions of momentum, it is simpler in the case of the oscillator to start directly from Schrödinger's equation in the momentum representation. Substituting in (23.1) the coordinate operator (15.12), we obtain the Hamiltonian in the p representation,

$$\widehat{H} = \frac{p^2}{2m} - \frac{m\omega^2\hbar}{2}\frac{d^2}{dp^2}.$$

The corresponding Schrödinger's equation $\widehat{H}a(p) = Ea(p)$ for the wave function $a(p)$ in the momentum representation is

$$\frac{d^2a(p)}{dp^2} + \frac{2}{m\omega^2\hbar^2}\left(E - \frac{p^2}{2m}\right)a(p) = 0.$$

This equation is of exactly the same form as (23.6); hence its solutions can be written down at once by analogy with (23.12). Thus we find the required probability distribution to be

$$|a(p)|^2\frac{dp}{2\pi\hbar} = \frac{1}{2^n n!\sqrt{\pi m\omega\hbar}}\exp\left(-\frac{p^2}{m\omega\hbar}\right)H_n^2\left(\frac{p}{\sqrt{m\omega\hbar}}\right).$$

2. Determine the lower limit of the possible values of the energy of an oscillator, using the uncertainty relation (16.7).

SOLUTION. Since $\overline{x^2} = \overline{x}^2 + (\delta x)^2, \overline{p^2} = \overline{p}^2 + (\delta p)^2$, (16.7) gives for the mean value of the energy of the oscillator

$$\overline{E} = \frac{m\omega^2}{2}\overline{x^2} + \frac{\overline{p^2}}{2m} \geqslant \frac{m\omega^2}{2}\overline{(\delta x)^2} + \frac{\overline{(\delta p)^2}}{2m} \geqslant \frac{m\omega^2\hbar^2}{8(\delta p)^2} + \frac{(\delta p)^2}{2m}.$$

On determining the minimum value of this expression (regarded as a function of δp), we find the lower limit of the mean values of the energy, and therefore that of all possible values: $E \geqslant \hbar\omega/2$.

3. Find the wave functions of the states of a linear oscillator that minimize the uncertainty relation, i.e. in which the standard deviations of the coordinate and momentum in the wave packet are related by $\delta p \delta x = \hbar/2$ (E. Schrödinger 1926).[15]

SOLUTION. The required wave functions must have the form

$$\Psi(x,t) = \frac{1}{(2\pi)^{1/4}(\delta x)^{1/2}} \exp\left\{ i\frac{\overline{p}x}{\hbar} - \frac{(x-\overline{x})^2}{4(\delta x)^2} - i\varphi(t) \right\}. \tag{1}$$

Their dependence on the coordinate at any instant is in accordance with (16.8), $\overline{x} = \overline{x}(t)$ and $\overline{p} = \overline{p}(t) = m\dot{\overline{x}}(t)$ being the mean values of the coordinate and the momentum; according to (19.3), we have, for a linear oscillator ($U = m\omega^2 x^2/2$), $\widehat{\dot{p}} = -m\omega^2 x$, and therefore for the mean value $\dot{\overline{p}} = -m\omega^2\overline{x}$ or

$$\ddot{\overline{x}} + \omega^2\overline{x} = 0 \tag{2}$$

i.e. the function satisfies the classical equation of motion. The constant factor in (1) is determined by the normalization condition $\int_{-\infty}^{+\infty} |\Psi|^2 dx = 1$ in addition to this factor, Ψ may contain a phase factor with a time-dependent phase $\varphi(t)$. The unknown constant δx and the unknown function $\varphi(t)$ are found by substituting (1) in the wave equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + \frac{m\omega^2 x^2}{2}\Psi = i\hbar\frac{\partial\Psi}{\partial t}.$$

With (2), the substitution gives

$$\left(\frac{x^2}{2} - x\overline{x}\right)\left(\frac{m^2\omega^2}{\hbar^2} - \frac{1}{4(\delta x)^4}\right) + \left[\frac{m^2\dot{\overline{x}}^2}{2\hbar^2} - \frac{\overline{x}^2}{8(\delta x)^4} + \frac{1}{4(\delta x)^2} - \frac{m}{\hbar}\dot{\varphi}(t)\right] = 0.$$

Hence $(\delta x)^2 = \hbar/(2m)$ and

$$\dot{\varphi} = \frac{m}{2\hbar}(\dot{\overline{x}}^2 - \omega^2\overline{x}^2) + \frac{\omega}{2}, \quad \varphi = \frac{1}{2\hbar}\overline{p}x + \frac{\omega}{2}t.$$

---

[15] These are called *coherent states.*

Thus we have finally

$$\Psi(x,t) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{\frac{\mathrm{i}\overline{p}x}{\hbar} - \frac{m\omega(x-\overline{x})^2}{2\hbar}\right\} \exp\left\{-\frac{\mathrm{i}\omega t}{2} - \frac{\mathrm{i}\overline{p}\overline{x}}{2\hbar}\right\} \qquad (3)$$

When $\overline{x} = 0$ and $\overline{p} = 0$, this becomes $\psi_0(x)\mathrm{e}^{\mathrm{i}\omega t/2}$, the wave function of the oscillator ground state.

The mean energy of the oscillator in a coherent state is

$$\overline{E} = \frac{\overline{p^2}}{2m} + \frac{m\omega^2\overline{x^2}}{2} = \frac{\overline{p}^2}{2m} + \frac{m\omega^2\overline{x}^2}{2} + \frac{\hbar\omega}{2} \equiv \hbar\omega\left(\overline{n} + \frac{1}{2}\right); \qquad (4)$$

the quantity $\overline{n}$ is the mean "number" of quanta $\hbar\omega$ in the state. We see that the coherent state is completely specified by the function $\overline{x}(t)$ satisfying the classical equation (2). The general form of this function may be given as

$$\frac{m\omega\overline{x} + \mathrm{i}\overline{p}}{\sqrt{2m\hbar\omega}} a\mathrm{e}^{\mathrm{i}\omega t}, \quad |a|^2 = \overline{n}. \qquad (5)$$

The function (3) can be expanded in wave functions of the stationary states of the oscillator:

$$\Psi = \sum_{n=0}^{\infty} a_n\Psi_n, \quad \Psi_n(x,t) = \psi_n(x)\exp\left\{-\mathrm{i}\left(n + \frac{1}{2}\right)\omega t\right\}.$$

The coefficients in this expansion are (cf. §41, Problem 1)

$$a_n = \int_{-\infty}^{+\infty} \Psi_n^* \Psi \mathrm{d}x. \qquad (6)$$

The probability for the oscillator to be in the nth state is therefore

$$\omega_n = |a_n|^2 = \mathrm{e}^{-\overline{n}}\frac{\overline{n}^n}{n!} \qquad (7)$$

the Poisson distribution.

4. Determine the energy levels for a particle moving in a field of potential energy (Fig. 3; P. M. Morse)

$$U(x) = A\left(\mathrm{e}^{-2\alpha x} - 2\mathrm{e}^{-\alpha x}\right).$$

SOLUTION. The spectrum of positive eigenvalues of the energy is continuous (and the levels are not degenerate), while the spectrum of negative eigenvalues is discrete.

Schrödinger's equation reads

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}\left(E - A\mathrm{e}^{-2\alpha x} + 2A\mathrm{e}^{-\alpha x}\right)\psi = 0.$$

We introduce a new variable

$$\xi = \frac{2\sqrt{2mA}}{\alpha\hbar}e^{-\alpha x}$$

(taking values from 0 to ∞) and the notation (we consider the discrete spectrum, so that $E < 0$)

$$s = \frac{\sqrt{-2mE}}{\alpha\hbar}, \quad n = \frac{\sqrt{2mA}}{\alpha\hbar} + \left(s + \frac{1}{2}\right) \tag{1}$$

Schrödinger's equation then takes the form

$$\psi'' + \frac{1}{\xi}\psi' + \left(-\frac{1}{4} + \frac{n+s+1/2}{\xi} - \frac{s^2}{\xi^2}\right)\psi = 0$$

As $\xi \to \infty$, the function ψ behaves asymptotically as $e^{\pm\xi/2}$, while as $\xi \to 0$ it is proportional to $\xi\pm s$. From considerations of finiteness we must choose the solution which behaves as $e^{-\xi/2}$ as $\xi \to \infty$ and as ξ^s as $\xi \to 0$. We make the substitution

$$\psi = e^{-\xi/2}\xi^s\omega(\xi)$$

and obtain for ω the equation

$$\xi\omega'' + (2s + 1 - \xi)\omega' + n\omega = 0, \tag{2}$$

which has to be solved with the conditions that ω is finite as $\xi \to 0$, while as $\xi \to \infty$, ω tends to infinity not more rapidly than every finite power of ξ. Equation (2) is the equation for a confluent hypergeometric function (see §d of the Mathematical Appendices):
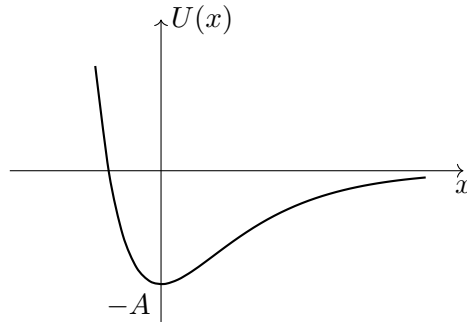
$$\omega = F(-n, 2s + 1, \xi).$$

FIG. 3

A solution satisfying the required conditions is obtained for non-negative integral n (when the function F reduces to a polynomial). According to the definitions (1), we thus obtain for the energy levels the values

$$-E_n = \left[1 - \frac{\alpha\hbar}{\sqrt{2mA}}\left(n + \frac{1}{2}\right)\right]^2.$$

where n takes positive integral values from zero to the greatest value for which (so that the parameter s is positive in accordance with its definition). Thus the discrete spectrum contains only a limited number of levels. If

$$\frac{\sqrt{2mA}}{\alpha\hbar} < \frac{1}{2},$$

there is no discrete spectrum at all.

5. The same as Problem 4, but with $U = -\dfrac{U_0}{\cosh^2 \alpha x}$ (Fig. 4).

SOLUTION. The spectrum of positive eigenvalues of the energy is continuous, while that of negative values is discrete; we shall consider the latter. Schrödinger's equation is

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar}\left(E + \frac{U_0}{\cosh^2 \alpha x}\right)\psi = 0.$$

We put $\xi = \tanh \alpha x$ and use the notation

$$\varepsilon = \frac{\sqrt{-2mE}}{\hbar\alpha}, \quad \frac{2mU_0}{\alpha^2\hbar^2} = s(s+1), \quad s = \frac{1}{2}\left(-1 + \sqrt{1 + \frac{8mU_0}{\alpha^2\hbar^2}}\right),$$

obtaining

$$\frac{d}{d\xi}\left[(1-\xi^2)\frac{d\psi}{d\xi}\right] + \left[s(s+1) - \frac{\varepsilon^2}{1-\xi^2}\right]\psi = 0.$$

This is the equation of the associated Legendre polynomials; it can be brought to hypergeometric form by making the substitution

FIG. 4

$$\psi = (1\xi^2)^{\varepsilon/2}w(\xi)$$

and temporarily changing the variable to $\frac{1}{2}(1) = u$:

$$u(1-u)w'' + (\varepsilon+1)(1-2u)w' - (\varepsilon-s)(\varepsilon+s+1)w = 0.$$

The solution finite for $\xi = 1$ (i.e. for $x = \infty$) is

$$\psi = (1\xi^2)^{\varepsilon/2}w(\xi)F[\varepsilon - s, \varepsilon + s + 1, \varepsilon - 1, (1-\xi)/2].$$

If ψ remains finite for $\xi = 1$ (i.e. for $x = -\inf$), we must have $\varepsilon - s = -n$ where $n = 0, 1, 2, \ldots$; then F is a polynomial of degree n, which is finite for $\xi = 1$.

Thus the energy levels are determined by $\varepsilon - s = -n$, or

$$E_n = -\frac{\hbar^2\alpha^2}{8m}\left[-(2n+1) + \sqrt{1 + \frac{8mU_0}{\alpha^2\hbar^2}}\right]^2.$$

There is a finite number of levels, determined by the condition $\varepsilon > 0$, i.e. $n < s$.

§ 24. Motion in a homogeneous field

Let us consider the motion of a particle in a homogeneous external field. We take the direction of the field as the axis of x; let F be the force acting
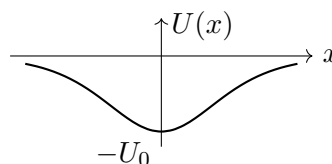
on the particle in this field. In an electric field of intensity E, this force is $F = eE$, where e is the charge on the particle.

The potential energy of the particle in the homogeneous field is of the form $U = Fx + \text{const}$; choosing the constant so that $U = 0$ for $x = 0$, we have $U = -Fx$. Schrödinger's equation for this problem is

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{2m}{\hbar}(E + Fx)\psi = 0. \tag{24.1}$$

Since U tends to $+\infty$ as $x \to -\infty$, and vice versa, it is clear that the energy levels form a continuous spectrum occupying the whole range of energy values E from $-\infty$ to $+\infty$. None of these eigenvalues is degenerate, and they correspond to motion which is finite towards $x = -\infty$ and infinite towards $x = +\infty$.

Instead of the coordinate x, we introduce the dimensionless variable

$$\xi = \left(x + \frac{E}{F}\right)\left(\frac{2mF}{\hbar}\right)^{1/3}. \tag{24.2}$$

Equation (24.1) then takes the form

$$\psi'' + \xi\psi = 0. \tag{24.3}$$

This equation does not contain the energy parameter. Hence, if we obtain a solution of it which satisfies the necessary conditions of finiteness, we at once have the eigenfunction for arbitrary values of the energy.

The solution of equation (24.3) which is finite for all x has the form (see §b of the Mathematical Appendices)

$$\psi(\xi) = A\Phi(-\xi), \tag{24.4}$$

where

$$\Phi(\xi) = \frac{1}{\sqrt{\pi}} \int_0^\infty \cos\left(\frac{1}{3}u^3 + u\xi\right) \mathrm{d}u$$

is called the *Airy function*, while A is a normalization factor which we shall determine below.

As $\xi \to -\infty$, the function $\psi(\xi)$ tends exponentially to zero. The asymptotic expression which determines $\psi(\xi)$ for large negative values of ξ is (see (b.4))

$$\psi(\xi) \approx \frac{A}{2|\xi|^{1/4}} \exp\left(-\frac{2}{3}|\xi|^{3/2}\right) \tag{24.5}$$

For large positive values of ξ, the asymptotic expression for $\psi(\xi)$ is (see (b.5))[16]

$$\psi(\xi) = \frac{A}{\xi^{1/4}} \sin\left(\frac{2}{3}\xi^{3/2} + \frac{\pi}{4}\right). \tag{24.6}$$

Using the general rule (5.4) for the normalization of eigenfunctions of a continuous spectrum, let us reduce the function (24.4) to the form normalized by the delta function of energy, for which

$$\int_{-\infty}^{+\infty} \psi(\xi)\psi(\xi')\mathrm{d}x\delta(E'-E). \tag{24.7}$$

In §21 we gave a simple method of determining the normalization coefficient by means of the asymptotic expression for the wave functions. Following this method, we represent the function (24.6) as the sum of two travelling waves:

$$\psi(\xi) \approx \frac{A}{2\xi^{1/4}}\left\{\exp\left[\mathrm{i}\left(\frac{2}{3}\xi^{3/2} - \frac{\pi}{4}\right)\right] + \exp\left[-\mathrm{i}\left(\frac{2}{3}\xi^{3/2} - \frac{\pi}{4}\right)\right]\right\}.$$

The current density, calculated from each of these two terms, is

$$v\left(\frac{A}{2\xi^{1/4}}\right)^2 = \sqrt{\frac{2}{m}(E+Fx)}\left(\frac{A}{2\xi^{1/4}}\right)^2 = A^2\frac{(2\hbar F)^{1/3}}{4m^{2/3}}.$$

and equating this to $1/2\pi\hbar$ we find

$$A = \frac{(2m)^{1/3}}{\pi^{1/2}F^{1/6}\hbar^{2/3}}. \tag{24.8}$$

PROBLEM

Determine the wave functions in the momentum representation for a particle in a homogeneous field.

SOLUTION. The Hamiltonian in the momentum representation is

$$\widehat{H} = \frac{p^2}{2m} - \mathrm{i}\hbar F\frac{\mathrm{d}}{\mathrm{d}p},$$

so that Schrödinger's equation for the wave function $a(p)$ has the form

$$-\mathrm{i}\hbar F\frac{\mathrm{d}a}{\mathrm{d}p} + \left(\frac{p^2}{2m} - E\right)a.$$

---

[16]) It may be noted, by way of anticipation, that the asymptotic expressions (24.5) and (24.6) correspond to the quasi-classical expressions for the wave function in the classically inaccessible and accessible regions (§47).

Solving this equation, we find the required functions

$$a_E(p) = \frac{1}{\sqrt{2\pi\hbar F}} \exp\left\{ \frac{\mathrm{i}}{\hbar F} F\left(Ep - \frac{p^3}{6m} \right) \right\}.$$

These functions are normalized by the condition

$$\int_{-\infty}^{+\infty} a_E^*(p) a_{E'}(p) \mathrm{d}p = \delta(E' - E).$$

§ 25. The transmission coefficient

Let us consider the motion of particles in a field of the type shown in Fig. 5: $U(x)$ increases monotonically from one constant limit ($U = 0$ as $x \to -\infty$) to another ($U = U_0$ as $x \to +\infty$). According to classical mechanics, a particle of energy $E < U_0$ moving in such a field from left to right, on reaching such a "potential wall", is reflected from it, and begins to move in the opposite direction; if, however, $E > U_0$, the particle continues to move in its original direction, though with diminished velocity. In quantum mechanics, a new phenomenon appears: even for $E > U_0$, the particle may be reflected from the potential wall. The probability of reflection must in principle be calculated as follows.
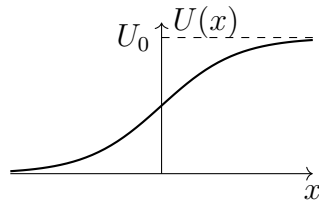
FIG. 5

Let the particle be moving from left to right. For large positive values of x, the wave function must describe a particle which has passed "above the wall" and is moving in the positive direction of x, i.e. it must have the asymptotic form

$$\text{for } x \to \infty: \quad \psi \approx A\mathrm{e}^{\mathrm{i}k_2 x}, \quad k_2 = \frac{1}{\hbar}\sqrt{2m(E - U_0)} \qquad (25.1)$$

and A is a constant. To find the solution of Schrödinger's equation which satisfies this boundary condition, we calculate the asymptotic expression for $x \to -\infty$; it is a linear combination of the two solutions of the equation of free motion, i.e. it has the form

$$\text{for } x \to -\infty: \quad \psi \approx \mathrm{e}^{\mathrm{i}k_1 x} + B\mathrm{e}^{-k_1 x}, \quad k_1 = \frac{1}{\hbar}\sqrt{2mE}. \qquad (25.2)$$

The first term corresponds to a particle incident on the wall (we suppose ψ normalized so that the coefficient of this term is unity); the second term represents a particle reflected from the wall. The probability current density

in the incident wave is k_1, in the reflected wave $k_1|B|^2$, and in the transmitted wave $k_2|A|^2$. We define the *transmission coefficient D* of the particle as the ratio of the probability current density in the transmitted wave to that in the incident wave:

$$D = \frac{k_2}{k_1}|A|^2. \tag{25.3}$$

Similarly we can define the reflection coefficient R as the ratio of the density in the reflected wave to that in the incident wave. Evidently $R = 1D$:

$$R = |B|^2 = 1 - \frac{k_2}{k_1}|A|^2 \tag{25.4}$$

(this relation between A and B is automatically satisfied).

If the particle moves from left to right with energy $E < U_0$, then k_2 is purely imaginary, and the wave function decreases exponentially as $x \to +\infty$. The reflected current is equal to the incident one, i.e. we have "total reflection" of the particle from the potential wall. We emphasize, however, that in this case the probability of finding the particle in the region where $E < U$ is still different from zero, though it diminishes rapidly as x increases.

In the general case of an arbitrary stationary state (with energy $E > U_0$), the asymptotic form of the wave function is given, both for $x \to -\infty$ and for $x \to +\infty$, by a sum of waves propagated in each direction:

$$\begin{aligned}
\psi &= A_1 \mathrm{e}^{ik_1x} + B_1 \mathrm{e}^{-ik_1x} \quad \text{for} \quad x \to -\infty, \\
\psi &= A_2 \mathrm{e}^{ik_2x} + B_2 \mathrm{e}^{-ik_2x} \quad \text{for} \quad x \to +\infty.
\end{aligned} \tag{25.5}$$

Since these expressions are asymptotic forms of the same solution of a linear differential equation, there must be a linear relation between the coefficients A_1, B_1 and A_2, B_2. Let $A_2 = \alpha A_1 + \beta B_1$, where α, β are constants (in general complex) which depend on the specific form of the field $U(x)$. The corresponding relation for B_2 can then be written down from the fact that Schrödinger's equation is real. This shows that, if ψ is a solution of a given Schrödinger's equation, the complex conjugate function ψ^* is also a solution. The asymptotic forms

$$\begin{aligned}
\psi^* &= A_1^* \mathrm{e}^{-ik_1x} + B_1^* \mathrm{e}^{ik_1x} \quad \text{for} \quad x \to -\infty, \\
\psi^* &= A_2^* \mathrm{e}^{-ik_2x} + B_2^* \mathrm{e}^{ik_2x} \quad \text{for} \quad x \to +\infty
\end{aligned}$$

differ from (25.5) only in the nomenclature of the constant coefficients; we therefore have $B_2^* = \alpha B_1^* + \beta A_1^*$ or $B2 = \alpha^* B_1 + \beta^* A_1$. Thus the coefficients in (25.5) are related by equations of the form

$$A_2 = \alpha A_1 + \beta B_1, \quad B_2 = \beta^* A_1 + \alpha^* B_1. \tag{25.6}$$

The condition of constant current along the x-axis leads to the relation

$$k_1(|A_1|^2 - |B_1|^2) = k_2(|A_2|^2 - |B_2|^2).$$

Expressing A_2, B_2 in terms of A_1, B_1 by (25.6), we find

$$|\alpha|^2 - |\beta|^2 = \frac{k_1}{k_2}. \tag{25.7}$$

Using the relation (25.6), we can show, in particular, that the reflection coefficients are equal (for a given energy $E > U_0$) for particles moving in the positive and negative directions of the x-axis; the former case corresponds to putting $B_2 = 0$ in (25.5), and the latter case to $A_1 = 0$. In these two cases, $B_1/A_1 = \beta^*/\alpha^*$ and $A_2/B_2 = \beta/\alpha^*$ respectively. The corresponding reflection coefficients are

$$R_1 = \left|\frac{B_1}{A_1}\right|^2 = \left|\frac{\beta^*}{\alpha^*}\right|^2, \quad R_2 = \left|\frac{A_2}{B_2}\right|^2 = \left|\frac{\beta}{\alpha^*}\right|^2$$

whence it is clear that $R_1 = R_2$.

It is natural to call $B_1/A_1 = -\beta^*/\alpha^*$ and $A_2/B_2 = \beta/\alpha^*$ the *reflection amplitudes* for motion in the positive and negative directions respectively. They are equal in modulus but may have different phase factors.

PROBLEMS

1. Determine the reflection coefficient of a particle from a rectangular potential wall (Fig. 6); the energy of the particle $E > U_0$.
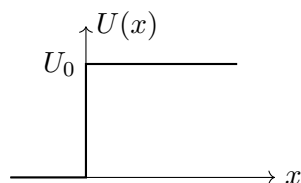
SOLUTION. Throughout the region $x > 0$, the wave function has the form (25.1), while in the region $x < 0$ its form is (25.2). The constants A and B are determined from the condition that ψ and $\mathrm{d}\psi/\mathrm{d}x$ are continuous at $x = 0$:

$$1 + B = A, \quad k_1(1 - B) = k_2 A,$$

whence

$$A = \frac{2k_1}{k_1 + k_2}, \quad B = \frac{k_1 - k_2}{k_1 + k_2}.$$

FIG. 6

The reflection coefficient[17] is (25.4)

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2 \left(\frac{p_1 - p_2}{p_1 + p_2}\right)^2$$

---

[17] In the limiting case of classical mechanics, the reflection coefficient must become zero. The expression obtained here, however, does not contain the quantum constant at all. This apparent contradiction is explained as follows. The classical limiting case is that in which the de Broglie wavelength of the particle $\lambda \sim \hbar/p$ is small in comparison with the characteristic dimensions of the problem, i.e. the distances over which the field $U(x)$ changes noticeably. In the schematic example considered, however, this distance is zero (at the point $x = 0$), so that the passage to the limit cannot be effected.

For $E = U_0$ ($k_2 = 0$), R becomes unity, while for $E \to \infty$ it tends to zero as $R = (U_0/4E)^2$.

2. Determine the transmission coefficient for a rectangular potential barrier (Fig. 7).

SOLUTION. Let E be greater than U_0, and suppose that the incident particle is moving from left to right. Then we have for the wave function in the different regions expressions of the form

$$\psi = \mathrm{e}^{\mathrm{i}k_1 x} + A\mathrm{e}^{-\mathrm{i}k_1 x} \qquad\qquad \text{for } x < 0,$$
$$\psi = B\mathrm{e}^{\mathrm{i}k_2 x} + B'\mathrm{e}^{-\mathrm{i}k_2 x} \qquad\qquad \text{for } 0 < x < a,$$
$$\psi = C\mathrm{e}^{\mathrm{i}k_1 x} \qquad\qquad \text{for } x > a.$$

(on the side $x > a$ there can be only the transmitted wave, propagated in the positive direction of x). The constants A, B, B' and C are determined from the conditions of continuity of ψ and $\mathrm{d}\psi/\mathrm{d}x$ at the points $x = 0$ and a. The transmission coefficient is determined as $D = k_1|C|^2/k_1 = |C|^2$. On calculating this, we obtain

$$D = \frac{4k_1^2 k_2^2}{(k_1^2 - k_2^2)^2 \sin^2 ak_2 + 4k_1^2 k_2^2}$$
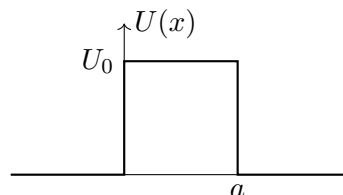

FIG. 7

For $E < U_0$, k_2 is a purely imaginary quantity; the corresponding expression for D is obtained by replacing k_2 by $\mathrm{i}\varkappa_2$ where $\hbar\varkappa_2 = \sqrt{2m(U_0 - E)}$:

$$D = \frac{4k_1^2 \varkappa_2^2}{(k_1^2 - \varkappa_2^2)^2 \sin^2 a\varkappa_2 + 4k_1^2 \varkappa_2^2}$$

3. Determine the reflection coefficient for a potential wall defined by the formula

$$U(x) = U_0/(1 + \mathrm{e}^{-\alpha x})$$

(Fig. 5); the energy of the particle is $E > U_0$.

SOLUTION. Schrödinger's equation is

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}\left(E - \frac{U_0}{1 + \mathrm{e}^{-\alpha x}}\right)\psi = 0.$$

We have to find a solution which, as $x \to +\infty$, has the form

$$\psi = \text{const} \cdot \mathrm{e}^{\mathrm{i}k_2}.$$

We introduce a new variable

$$\xi = -\mathrm{e}^{-\alpha x}.$$

(which takes values from $-\infty$ to 0), and seek a solution of the form

$$\psi = \xi^{-ik_2/\alpha}w(\xi),$$

where $w(\xi)$ tends to a constant as $\xi \to 0$ (i.e. as $x \to \infty$). For $w(\xi)$ we find an equation of hypergeometric type:

$$\xi(1-\xi)w'' + \left(1 - \frac{2i}{\alpha}k_2\right)(1-\xi)w' + \frac{1}{\alpha^2}(k_2^2 - k_1^2)w = 0,$$

which has as its solution the hypergeometric function

$$w = F\left[\frac{i}{\alpha}(k_1 - k_2), -\frac{i}{\alpha}(k_1 + k_2), -\frac{2i}{\alpha}k_2 + 1, \xi\right]$$

(we omit a constant factor). As $\xi \to 0$, this function tends to 1, i.e. it satisfies the condition imposed.

The asymptotic form of the function ψ as $\xi \to -\infty$ (i.e. $x \to -\infty$) is[18]

$$\psi \approx \xi^{-ik_2/\alpha}\left[C_1(-\xi)^{i(k_2-k_1)/\alpha} + C_2(-\xi)^{i(k_1+k_2)/\alpha}\right] =$$

$$= (-1)^{ik_2/\alpha}\left[C_1 e^{ik_1 x} + C_2 e^{-ik_2 x}\right],$$

where

$$C_1 = \frac{\Gamma(-(2i/\alpha)k_1)\Gamma(-(2i/\alpha)k_2 + 1)}{\Gamma(-(i/\alpha)(k_1 + k_2))\Gamma(-(i/\alpha)(k_1 + k_2) + 1)},$$

$$C_2 = \frac{\Gamma((2i/\alpha)k_1)\Gamma(-(2i/\alpha)k_2 + 1)}{\Gamma((i/\alpha)(k_1 - k_2))\Gamma((i/\alpha)(k_1 - k_2) + 1)}.$$

The required reflection coefficient is $R = |C_2/C_1|^2$ on calculating it by means of the well known formula

$$\Gamma(x)\Gamma(1 - x) = \frac{\pi}{\sin \pi x},$$

we have

$$R = \left\{\frac{\sinh[(\pi/\alpha)(k_1 - k_2)]}{\sinh[(\pi/\alpha)(k_1 + k_2)]}\right\}^2.$$

For $E = U_0$ ($k_2 = 0$), R becomes unity, while for $E \to \infty$ it tends to zero as

$$R = \left(\frac{\pi U_0}{\alpha\hbar}\right)^2 \frac{2m}{E} \exp\left(-\frac{4\pi}{\alpha\hbar}\sqrt{2mE}\right).$$

In the limiting case of classical mechanics. R becomes zero, as it should.

---

[18] See formula (e.6), in each of whose two terms we must take only the first term of the expansion, i.e. replace the hypergeometric functions of $1/z$ by unity.
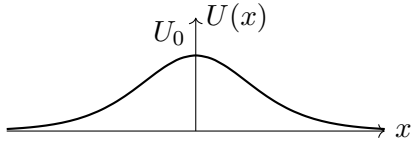
$U_0 \uparrow U(x)$

FIG. 8

4. Determine the transmission coefficient for a potential barrier defined by the formula

$$U(x) = \frac{U_0}{\cosh^2 \alpha x}$$

(Fig. 8); the energy of the particle is $E < U_0$.

SOLUTION. The Schrödinger's equation is the same as that obtained in the solution of Problem 5, §23; it is necessary merely to alter the sign of U_0 and to regard the energy E now as positive. A similar calculation gives the solution

$$\psi = (1 - \xi^2)^{-\frac{ik}{2\alpha}} F\left(-\frac{ik}{\alpha} - s, -\frac{ik}{\alpha} + s + 1, -\frac{ik}{\alpha} + 1, \frac{1-\xi}{2}\right), \qquad (1)$$

where

$$\xi = \tanh \alpha x, \quad k = \frac{1}{\hbar}\sqrt{2mE}, \quad s = \frac{1}{2}\left(-1 + \sqrt{1 - \frac{8mU_0}{\alpha^2 \hbar^2}}\right).$$

This solution satisfies the condition that, as $x \to \infty$ (i.e. as $\xi \to 1$, $(1-\xi) \approx 2e^{-2\alpha x}$), the wave function should include only the transmitted wave ($\sim e^{ikx}$). The asymptotic form of the wave function as $x \to -\infty (\xi \to 1)$ is found by transforming the hypergeometric function with the aid of formula (e.7):

$$\psi \sim e^{ikx}\frac{\Gamma(ik/\alpha)\Gamma(1 - ik/\alpha)}{\Gamma(-s)\Gamma(1 + s)} + e^{ikx}\frac{\Gamma(-ik/\alpha)\Gamma(1 - ik/\alpha)}{\Gamma(-ik/\alpha - s)\Gamma(-ik/\alpha + s + 1)}. \qquad (2)$$

Taking the squared modulus of the ratio of coefficients in this function, we obtain the following expression for the transmission coefficient $D = 1R$:

$$D = \frac{\sinh^2 \frac{\pi k}{\alpha}}{\sinh^2 \frac{\pi k}{\alpha} + \cos^2\left(\frac{\pi}{2}\sqrt{1 - \frac{8mU_0}{\hbar^2 \alpha^2}}\right)} \qquad \text{if} \quad \frac{8mU_0}{\hbar^2 \alpha^2} < 1,$$

$$D = \frac{\sinh^2 \frac{\pi k}{\alpha}}{\sinh^2 \frac{\pi k}{\alpha} + \cosh^2\left(\frac{\pi}{2}\sqrt{\frac{8mU_0}{\hbar^2 \alpha^2} - 1}\right)} \qquad \text{if} \quad \frac{8mU_0}{\hbar^2 \alpha^2} > 1.$$

The first of these formulae holds also for the case $U_0 < 0$, i.e. when the particle is passing over a potential well instead of a potential barrier. It is interesting to note that in that case $D = 1$ if $1 + 8m|U_0|/\hbar^2\alpha^2 = (2n + 1)^2$; thus, for certain values of the depth $|U_0|$ of the well, particles passing over it are not reflected. This is evident from equation (2), where the term in e^{ikx} vanishes for positive integral s.

5. Determine how the transmission coefficient tends to zero as $E \to 0$, assuming that the potential energy $U(x)$ decreases rapidly at distances $|x| \gg a$, where a is the dimension of the interaction region.

SOLUTION. For distances $k|x| \ll 1$, E can be neglected in Schrödinger's equation. If also $|x| \gg a$, the potential energy can also be neglected, and the equation becomes

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = 0,$$

the solution of this may be written as

$$\psi = a_1 + b_1 x, \quad x < 0, \qquad \psi = a_2 + b_2 x, \quad x > 0. \tag{1}$$

The relation between a_1, b_1 and a_2, b_2 can be found by solving the equation at distances $|x| \sim a$. It is linear:

$$a_1 = \rho a_2 + \mu b_2, \quad b_1 = \nu a_1 + \tau b_2. \tag{2}$$

The coefficients ρ, μ, ν and τ are real and independent of the energy, which does not appear in the equation.[19] The solution (1) must be the same as the first two terms in the expansion of (25.1) and (25.2) in powers of x, so that

$$a_1 = B + 1, \quad b_1 = \mathrm{i}k(1 - B), \quad a_2 = A, \quad b_2 = \mathrm{i}kA.$$

Substituting these in (2) and solving for A, we get, for small k: $A \approx 2\mathrm{i}k/\nu$, whence

$$D \approx \frac{4k^2}{\nu^2} \sim E$$

The transmission coefficient thus tends to zero in proportion to the particle energy. This is of course true for the examples in Problems 2 and 4.

---

[19] Since the flux is constant, $\rho\tau - \mu\nu = 1$.

CHAPTER IV

ANGULAR MOMENTUM

§ 26. Angular momentum

IN §15, to derive the law of conservation of momentum, we have made use of the homogeneity of space relative to a closed system of particles. Besides its homogeneity, space has also the property of isotropy: all directions in it are equivalent. Hence the Hamiltonian of a closed system cannot change when the system rotates as a whole through an arbitrary angle about an arbitrary axis. It is sufficient to require the fulfilment of this condition for an infinitely small rotation.

Let $\delta\boldsymbol{\varphi}$ be the vector of an infinitely small rotation, equal in magnitude to the angle $\delta\varphi$ of the rotation and directed along the axis about which the rotation takes place. The changes $\delta\boldsymbol{r}_a$ (in the radius vectors $\boldsymbol{r}_a$ of the particles) in such a rotation are

$$\delta\boldsymbol{r}_a = \delta\boldsymbol{\varphi} \times \boldsymbol{r}_a.$$

An arbitrary function $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$ is thereby transformed into the function

$$\psi(\boldsymbol{r}_1 + \delta\boldsymbol{r}_1, \boldsymbol{r}_2 + \delta\boldsymbol{r}_2, \dots) = \psi(\boldsymbol{r}_1 + \boldsymbol{r}_2 + \dots) + \sum_a \delta\boldsymbol{r}_a \cdot \nabla_a \psi =$$

$$= \psi(\boldsymbol{r}_1 + \boldsymbol{r}_2 + \dots) + \sum_a \delta\boldsymbol{\varphi} \times \boldsymbol{r}_a \cdot \nabla_a \psi =$$

$$= \left(1 + \delta\boldsymbol{\varphi} \cdot \sum_a \boldsymbol{r}_a \times \nabla_a \right) \psi(\boldsymbol{r}_1 + \boldsymbol{r}_2 + \dots).$$

The expression

$$1 + \delta\boldsymbol{\varphi} \cdot \sum_a \boldsymbol{r}_a \times \nabla_a$$

is the operator of an infinitely small rotation. The fact that an infinitely small rotation does not alter the Hamiltonian of the system is expressed (cf.

97

§15) by the commutability of the "rotation operator" with the operator $\widehat{H}$. Since $\delta\boldsymbol{\varphi}$ is a constant vector, this condition reduces to the relation

$$\left(\sum_a \boldsymbol{r}_a \times \nabla_a\right) \widehat{H} - \widehat{H} \left(\sum_a \boldsymbol{r}_a \times \nabla_a\right) = 0, \qquad (26.1)$$

which expresses a certain law of conservation.

The quantity whose conservation for a closed system follows from the property of isotropy of space is the *angular momentum* of the system (cf. *Mechanics*, §9). Thus the operator $\sum \boldsymbol{r}_a \times \nabla_a$ must correspond exactly, apart from a constant factor, to the total angular momentum of the system, and each of the terms $\boldsymbol{r}_a \times \nabla_a$ of this sum corresponds to the angular momentum of an individual particle.

The coefficient of proportionality must be put equal to $-i\hbar$; then the expression for the angular momentum operator of a particle is $-i\hbar\boldsymbol{r} \times \nabla = \boldsymbol{r} \times \widehat{\boldsymbol{p}}$ and corresponds exactly to the classical expression $\boldsymbol{r} \times \boldsymbol{p}$. Henceforward we shall always use the angular momentum measured in units of $\hbar$. The angular momentum operator of a particle, so defined, will be denoted by $\widehat{\boldsymbol{l}}$, and that of the whole system by $\widehat{\boldsymbol{L}}$. Thus the angular momentum operator of a particle is

$$\hbar\widehat{\boldsymbol{l}} = \boldsymbol{r} \times \widehat{\boldsymbol{p}} = -i\hbar\boldsymbol{r} \times \nabla \qquad (26.2)$$

or, in components,

$$\hbar\widehat{l_x} = y\widehat{p}_z - z\widehat{p}_y, \quad \widehat{l_y} = z\widehat{p}_x - x\widehat{p}_z, \quad \widehat{l_z} = x\widehat{p}_y - y\widehat{p}_x.$$

For a system which is in an external field, the angular momentum is in general not conserved. However, it may still be conserved if the field has a certain symmetry. Thus, if the system is in a centrally symmetric field, all directions in space at the centre are equivalent, and hence the angular momentum about this centre will be conserved. Similarly, in an axially symmetric field, the component of angular momentum along the axis of symmetry is conserved. All these conservation laws holding in classical mechanics are valid in quantum mechanics also.

In a system where angular momentum is not conserved, it does not have definite values in the stationary states. In such cases the mean value of the angular momentum in a given stationary state is sometimes of interest. It is easily seen that, in any non-degenerate stationary state, the mean value of the angular momentum is zero. For, when the sign of the time is changed, the energy does not alter, and, since only one stationary state corresponds to a given energy level, it follows that when t is changed into $-t$ the state of the system must remain the same. This means that the mean values of

all quantities, and in particular that of the angular momentum, must remain unchanged. But when the sign of the time is changed, so is that of the angular momentum, and we have $\overline{\boldsymbol{L}} = -\overline{\boldsymbol{L}}$, whence it follows that $\overline{\boldsymbol{L}} = 0$. The same result can be obtained by starting from the mathematical definition of the mean value as being the integral of $\psi * \widehat{\boldsymbol{L}}\psi$. The wave functions of non-degenerate states are real (see the end of §18). Hence the expression

$$\overline{\boldsymbol{L}} = -\mathrm{i}\hbar \int \psi^* \left(\sum_a \boldsymbol{r}_a \times \nabla_a \right) \psi \mathrm{d}q$$

is purely imaginary, and since $\overline{\boldsymbol{L}}$ must, of course, be real, it is evident that $\overline{\boldsymbol{L}} = 0$.

Let us derive the rules for commutation of the angular momentum operators with those of coordinates and linear momenta. By means of the relations (16.2) we easily find

$$\begin{aligned}
&\{\widehat{l}_x, x\} = 0, &&\{\widehat{l}_x, y\} = \mathrm{i}z, &&\{\widehat{l}_x, z\} = -\mathrm{i}y, \\
&\{\widehat{l}_y, y\} = 0, &&\{\widehat{l}_y, z\} = \mathrm{i}x, &&\{\widehat{l}_y, x\} = -\mathrm{i}z, \\
&\{\widehat{l}_z, z\} = 0, &&\{\widehat{l}_z, x\} = \mathrm{i}y, &&\{\widehat{l}_z, y\} = -\mathrm{i}x.
\end{aligned} \tag{26.3}$$

For instance,

$$\widehat{l}_x y - y\widehat{l}_x = \frac{1}{\hbar}(y\widehat{p}_z - z\widehat{p}_y)y - y(y\widehat{p}_z - z\widehat{p}_y)\frac{1}{\hbar} = -\frac{z}{\hbar}\widehat{p}_y, y = \mathrm{i}z$$

All the relations (26.3) can be written in tensor form as follows:

$$\{\widehat{l}_i, x_k\} = \mathrm{i}e_{ikl}x_l \tag{26.4}$$

where e_{ikl} is the *antisymmetric unit tensor* of rank three,[1] and summation is implied over those suffixes which appear twice (called *dummy suffixes*).

---

[1] The antisymmetric unit tensor of rank three, e_{ikl} (also called the unit axial tensor), is defined as a tensor antisymmetric in all three suffixes, with $e_{123} = 1$. It is evident that, of its 27 components, only 6 are not zero, namely those in which the suffixes i, k, l form some permutation of $1, 2, 3$. Such a component is $+1$ if the permutation i, k, l is obtained from $1, 2, 3$ by an even number of transpositions of pairs of figures, and is -1 if the number of transpositions is odd. Clearly

$$e_{ikl}e_{ikm} = 2\delta_{lm}, \quad e_{ikl}e_{ikl} = 6.$$

The components of the vector $\boldsymbol{C} = \boldsymbol{A} \times \boldsymbol{B}$ which is the vector product of the two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ can be written by means of the tensor e_{ikl} in the form

$$C_i = e_{ikl}A_k B_l.$$

It is easily seen that a similar commutation rule holds for the angular momentum and linear momentum operators:

$$\{\widehat{l}_i, \widehat{p}_k\} = \mathrm{i}e_{ikl}\widehat{p}_l. \tag{26.5}$$

By means of these formulae, it is easy to find the rules for commutation of the operators with one another. We have

$$\hbar(\widehat{l}_x\widehat{l}_y - \widehat{l}_y\widehat{l}_x) = \widehat{l}_x(z\widehat{p}_x - x\widehat{p}_z) - (z\widehat{p}_x - x\widehat{p}_z)\widehat{l}_x =$$
$$= (\widehat{l}_x z - z\widehat{l}_x)\widehat{p}_x - x(\widehat{l}_x\widehat{p}_z - \widehat{p}_z\widehat{l}_x) = \mathrm{i}y\widehat{p}_x + \mathrm{i}x\widehat{p}_y = \mathrm{i}\hbar\widehat{l}_z.$$

Thus

$$\{\widehat{l}_y, \widehat{l}_z\} = \mathrm{i}\widehat{l}_x, \quad \{\widehat{l}_z, \widehat{l}_x\} = \mathrm{i}\widehat{l}_y, \quad \{\widehat{l}_x, \widehat{l}_y\} = \mathrm{i}\widehat{l}_z, \tag{26.6}$$

or

$$\{\widehat{l}_i, \widehat{l}_k\} = \mathrm{i}e_{ikl}\widehat{l}_l \tag{26.7}$$

Exactly the same relations hold for the operators of the total angular momentum of the system. For, since the angular momentum operators of different individual particles commute, we have, for instance,

$$\sum_a \widehat{l}_{ay} \sum_a \widehat{l}_{az} - \sum_a \widehat{l}_{az} \sum_a \widehat{l}_{ay} = \sum_a \left(\widehat{l}_{ay}\widehat{l}_{az} - \widehat{l}_{az}\widehat{l}_{ay} \right).$$

Thus

$$\{\widehat{L}_y, \widehat{L}_z\} = \mathrm{i}\widehat{L}_x, \quad \{\widehat{L}_z, \widehat{L}_x\} = \mathrm{i}\widehat{L}_y, \quad \{\widehat{L}_x, \widehat{L}_y\} = \mathrm{i}\widehat{L}_z. \tag{26.8}$$

The relations (26.8) show that the three components of the angular momentum cannot simultaneously have definite values (except in the case where all three components simultaneously vanish: see below). In this respect the angular momentum is fundamentally different from the linear momentum, whose three components are simultaneously measurable.

From the operators $\widehat{L}_x, \widehat{L}_y, \widehat{L}_z$ we can form the operator of the square of the modulus of the angular momentum vector, and which we denote by :

$$\widehat{\boldsymbol{L}}^2 = \widehat{L}_x^2 + \widehat{L}_y^2 + \widehat{L}_z^2. \tag{26.9}$$

This operator commutes with each of the operators :

$$\{\widehat{\boldsymbol{L}}^2, \widehat{L}_x\} = 0, \quad \{\widehat{\boldsymbol{L}}^2, \widehat{L}_y\} = 0, \quad \{\widehat{\boldsymbol{L}}^2, \widehat{L}_z\} = 0. \tag{26.10}$$

Using (26.8), we have

$$\{\widehat{L}_x^2, \widehat{L}_z\} = \widehat{L}_x\{\widehat{L}_x, \widehat{L}_z\} + \{\widehat{L}_x, \widehat{L}_z\}\widehat{L}_x = -\mathrm{i}\left(\widehat{L}_x\widehat{L}_y + \widehat{L}_y\widehat{L}_x\right)$$

$$\{\widehat{L}_y^2, \widehat{L}_z\} = \mathrm{i}\left(\widehat{L}_x\widehat{L}_y + \widehat{L}_y\widehat{L}_x\right)$$

$$\{\widehat{L}_z^2, \widehat{L}_z\} = 0$$

Adding these equations, we have $\{\widehat{\boldsymbol{L}}^2, \widehat{L}_z\} = 0$. Physically, the relations (26.10) mean that the square of the angular momentum, i.e. its modulus, can have a definite value at the same time as one of its components.

Instead of the operators $\widehat{L}_x, \widehat{L}_y$ it is often more convenient to use the complex combinations

$$\widehat{L}_+ = \widehat{L}_x + \mathrm{i}\widehat{L}_y, \quad \widehat{L}_- = \widehat{L}_x - \mathrm{i}\widehat{L}_y \tag{26.11}$$

It is easily verified by direct calculation using (26.8) that the following commutation rules hold:

$$\{\widehat{L}_+, \widehat{L}_-\} = 2\widehat{L}_z, \quad \{\widehat{L}_z, \widehat{L}_+\} = \widehat{L}_+, \quad \{\widehat{L}_z, \widehat{L}_-\} = -\widehat{L}_-, \tag{26.12}$$

and it is also not difficult to see that

$$\widehat{\boldsymbol{L}}^2 = \widehat{L}_+\widehat{L}_- + \widehat{L}_z^2 - \widehat{L}_z = \widehat{L}_-\widehat{L}_+ + \widehat{L}_z^2 + \widehat{L}_z. \tag{26.13}$$

Finally, we shall give some frequently used expressions for the angular momentum operator of a single particle in spherical polar coordinates. Defining the latter by means of the usual relations

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta,$$

we have after a simple calculation

$$\widehat{l}_z = -\mathrm{i}\frac{\partial}{\partial\varphi}, \tag{26.14}$$

$$\widehat{l}_\pm = \mathrm{e}^{\pm\mathrm{i}\varphi}\left(\pm\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\varphi}\right). \tag{26.15}$$

Substitution in (26.13) gives the squared angular momentum operator of the particle:

$$\widehat{l}^2 = -\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\varphi}\right)\right] \tag{26.16}$$

It should be noticed that this is, apart from a factor, the angular part of the Laplacian operator.

§ 27. Eigenvalues of the angular momentum

In order to determine the eigenvalues of the component, in some direction, of the angular momentum of a particle, it is convenient to use the expression for its operator in spherical polar coordinates, taking the direction in question as the polar axis. According to formula (26.14), the equation $\widehat{l}_z\psi = l_z\psi$ can be written in the form

$$-\mathrm{i}\frac{\partial\psi}{\partial\varphi} = l_z\psi. \tag{27.1}$$

Its solution is

$$\psi = f(r,\theta)\mathrm{e}^{\mathrm{i}l_z\varphi},$$

where $f(r,\theta)$ is an arbitrary function of r and θ. If the function ψ is to be single-valued, it must be periodic in φ, with period 2π. Hence we find[2]

$$l_z = m, \quad m = 0, \pm1, \pm2, \ldots \tag{27.2}$$

Thus the eigenvalues l_z are the positive and negative integers, including zero. The factor depending on φ, which characterizes the eigenfunctions of the operator $\widehat{l}_z$, is denoted by

$$\Psi_m(\varphi) = \frac{1}{\sqrt{2}}\mathrm{e}^{\mathrm{i}m\varphi}. \tag{27.3}$$

These functions are normalized so that

$$\int_0^{2\pi} \Psi_m^*(\varphi)\Psi_{m'}(\varphi)\mathrm{d}\varphi = \delta_{mm'}. \tag{27.4}$$

The eigenvalues of the z-component of the total angular momentum of the system are evidently also equal to the positive and negative integers:

$$L_z = M, \quad M = 0, \pm1, \pm2, \ldots \tag{27.5}$$

(this follows at once from the fact that the operator $\widehat{L}_z$ is equal to the sum of the commuting operators $\widehat{l}_z$ for the individual particles).

Since the direction of the z-axis is in no way distinctive, it is clear that the same result is obtained for $\widehat{L}_x$, $\widehat{L}_y$ and in general for the component of the angular momentum in any direction: they can all take integral values only. At first sight this result may appear paradoxical, particularly if we apply it to two directions infinitely close to each other. In fact, however, it must be

---

[2] The customary notation for the eigenvalues of the angular momentum component is m, which also denotes the mass of a particle, but this should not lead to any confusion.

remembered that the only common eigenfunction of the operators $\widehat{L}_x, \widehat{L}_y, \widehat{L}_z$ corresponds to the simultaneous values

$$L_x = L_y = L_z = 0;$$

in this case the angular momentum vector is zero, and consequently so is its projection upon any direction. If even one of the eigenvalues L_x, L_y, L_z is not zero, the operators $\widehat{L}_x, \widehat{L}_y, \widehat{L}_z$ have no common eigenfunctions. In other words, there is no state in which two or three of the angular momentum components in different directions simultaneously have definite values different from zero, so that we can say only that one of them is integral.

The stationary states of a system which differ only in the value of M have the same energy; this follows from general considerations, based on the fact that the direction of the z-axis is in no way distinctive. Thus the energy levels of a system whose angular momentum is conserved (and is not zero) are always degenerate.[3]

Let us now look for the eigenvalues of the square of the angular momentum. We shall show how these values may be found, starting from the commutation rules (26.8) only. We denote by ψ_M the wave functions of the stationary states with the same value of $\boldsymbol{L}^2$, belonging to one degenerate energy level, and distinguished by the value of M.[4]

First of all we note that, since the two directions of the z-axis are physically equivalent, for every possible positive value $M = |M|$ there is a corresponding negative value $M = |M|$. Let L (a positive integer or zero) denote the greatest possible value of $|M|$ for a given $\widehat{\boldsymbol{L}}^2$. The existence of this upper limit follows because is the operator $\widehat{\boldsymbol{L}}^2 - \widehat{L}_z^2 = \widehat{L}_x^2 + \widehat{L}_y^2$ of the essentially positive physical quantity $L_x^2 + L_y^2$, and its eigenvalues therefore cannot be negative.

---

[3] This is a particular case of the general theorem, mentioned in §10, which states that the levels are degenerate when two or more conserved quantities exist whose operators do not commute. Here the components of the angular momentum are such quantities.

[4] Here it is supposed that there is no additional degeneracy leading to the same value of the energy for different values of the squared angular momentum. This is true for a discrete spectrum (except for the case of what is called *accidental degeneracy* in a Coulomb field; see §36) and in general untrue for the energy levels of a continuous spectrum. However, even when such additional degeneracy is present, we can always choose the eigenfunctions so that they correspond to states with definite values of $\boldsymbol{L}^2$, and then we can choose from these the states with the same values of E and $\widehat{\boldsymbol{L}}^2$. This is mathematically expressed by the fact that the matrices of commuting operators can always be simultaneously brought into diagonal form. In what follows we shall, in such cases, speak, for the sake of brevity, as if there were no additional degeneracy, bearing in mind that the results obtained do not in fact depend on this assumption, by what we have just said.

Applying the operator $\widehat{L}_z\widehat{L}_\pm$ to the eigenfunction ψ_M of the operator $\widehat{L}_z$ and using the commutation rule 26.12, we obtain

$$\widehat{L}_z\widehat{L}_\pm\psi_M = (M+1)\widehat{L}_\pm\psi_M. \tag{27.6}$$

Hence we see that the function $\widehat{L}_\pm\psi_M$ is (apart from a normalization constant) the eigenfunction corresponding to the value $M \pm 1$ of the quantity L_z:

$$\psi_{M+1} = \text{const} \cdot \widehat{L}_+\psi_M, \quad \psi_{M-1} = \text{const} \cdot L_-\psi_M. \tag{27.7}$$

If we put $M = L$ in the first of these equations, we must have identically

$$\widehat{L}_+\psi_L = 0 \tag{27.8}$$

since there is by definition no state with $M > L$. Applying the operator $\widehat{L}_-$ to this equation and using the relation (26.13), we obtain

$$\widehat{L}_-\widehat{L}_+\psi_L = (\widehat{\boldsymbol{L}}^2 - \widehat{L}_z^2 - \widehat{L}_z)\psi_L = 0.$$

Since, however, the ψ_M are common eigenfunctions of the operators $\widehat{\boldsymbol{L}}^2$ and $\widehat{L}_z$, we have

$$\widehat{\boldsymbol{L}}^2\psi_L = \boldsymbol{L}^2\psi_L, \quad \widehat{L}_z^2\psi_L = L^2\psi_L, \quad \widehat{L}_z\psi_L = L\psi_L,$$

so that the equation found above gives

$$\boldsymbol{L}^2 = L(L+1). \tag{27.9}$$

Formula (27.9) determines the required eigenvalues of the square of the angular momentum; the number L takes all positive integral values, including zero. For a given value of L, the component $L_z = M$ of the angular momentum can take the values

$$M = L, L-1, \ldots, -L, \tag{27.10}$$

i.e. $2L+1$ different values in all. The energy level corresponding to the angular momentum L thus has $(2L+1)$-fold degeneracy; this is usually called degeneracy with respect to the direction of the angular momentum. A state with angular momentum $L = 0$ (when all three components are also zero) is not degenerate. The wave function of such a state is spherically symmetric, as is evident from the fact that the change in it under any infinitesimal rotation, given by $\widehat{\boldsymbol{L}}\psi$, is in this case zero.

We shall often, for the sake of brevity, and in accordance with custom, speak of the "angular momentum" L of a system, understanding by this an

angular momentum whose square is $L(L+1)$; the z-component of the angular momentum is usually called just the "angular momentum component".

The angular momentum of a single particle is denoted by the small letter l, for which formula (27.9) becomes

$$\boldsymbol{l}^2 = l(l+1). \tag{27.11}$$

Let us calculate the matrix elements of the quantities L_x and L_y in a representation in which L_z and $\boldsymbol{L}^2$, as well as the energy, are diagonal (M. Born, W. Heisenberg and P. Jordan 1926). First of all, we note that, since the operators $\widehat{L}_x$ and $\widehat{L}_y$ commute with the Hamiltonian, their matrices are diagonal with respect to the energy, i.e. all matrix elements for transitions between states of different energy (and different angular momentum L) are zero. Thus it is sufficient to consider the matrix elements for transitions within a group of states with different values of M, corresponding to a single degenerate energy level.

It is seen from formulae (27.7) that, in the matrices of the operators $\widehat{L}_+$ and $\widehat{L}_-$, only those elements are different from zero which correspond to transitions $M - 1 \to M$ and $M \to M - 1$ respectively. Taking this into account, we find the diagonal matrix elements on both sides of the equation (26.13), obtaining[5]

$$L(L + 1) = \langle M|L_+|M - 1\rangle\langle M - 1|L_-|M\rangle + M^2 - M.$$

Noticing that, since the operators and are Hermitian,

$$\langle M - 1|L_-|M\rangle = \langle M|L_+|M - 1\rangle^*,$$

we can rewrite this equation in the form

$$|\langle M|L_+|M - 1\rangle|^2 = L(L + 1) - M(M + 1) = (L - M + 1)(L + M),$$

whence[6]

$$\langle M|L_+|M - 1\rangle = \langle M - 1|L_-|M\rangle = \sqrt{(L + M)(L - M + 1)}. \tag{27.12}$$

Hence we have for the non-zero matrix elements of the quantities L_x and L_y themselves

$$\langle M|L_x|M - 1\rangle = \langle M - 1|L_x|M\rangle = \frac{1}{2}\sqrt{(L + M)(L - M + 1)},$$
$$\langle M|L_y|M - 1\rangle = -\langle M - 1|L_y|M\rangle = -\frac{i}{2}\sqrt{(L + M)(L - M + 1)}. \tag{27.13}$$

_____

[5] In the symbols for the matrix elements, we omit for brevity all suffixes with respect to which they are diagonal (including L).

[6] The choice of sign in this formula corresponds to the choice of the phase factors in the eigenfunctions of the angular momentum.

The diagonal elements in the matrices of the quantities L_x and L_y are zero. Since a diagonal matrix element gives the mean value of the quantity in the corresponding state, it follows that the mean values $\overline{L_x}$ and $\overline{L_y}$ are zero in states having definite values of L_z. Thus, if the angular-momentum component in a given direction in space has a definite value, the vector $\overline{\boldsymbol{L}}$ itself is in that direction.

§ 28. Eigenfunctions of the angular momentum

The wave function of a particle is not completely determined when the values of l and m are prescribed. This is seen from the fact that the expressions for the operators of these quantities in spherical polar coordinates contain only the angles θ and φ, so that their eigenfunctions can contain an arbitrary factor depending on r. We shall here consider only the angular part of the wave function which characterizes the eigenfunctions of the angular momentum, and denote this by $Y_{lm}(\theta, \varphi)$, with the normalization condition:

$$\int |Y_{lm}|^2 \mathrm{d}o = 1,$$

where $\mathrm{d}o = \sin\theta \mathrm{d}\theta \mathrm{d}\varphi$ is an element of solid angle.

We shall see that the problem of determining the common eigenfunctions of the operators $\widehat{l}^2$ and $\widehat{l}_z$ admits of separation of the variables θ and φ, and these functions can be sought in the form

$$Y_{lm} = \Psi_m(\varphi)\Theta_{lm}(\theta), \tag{28.1}$$

where $\Theta_m(\varphi)$ are the eigenfunctions of the operator $\widehat{l}_z$, which are given by formula (27.3). Since the functions Θ_m are already normalized by the condition (27.4), the Θ_{lm} must be normalized by the condition

$$\int_0^\pi |\Theta_{lm}|^2 \sin\theta \mathrm{d}\theta. \tag{28.2}$$

The functions Y_{lm} with different l or m are automatically orthogonal:

$$\int_0^{2\pi} \int_0^\pi Y_{l'm'}^* Y_{lm} \sin\theta \mathrm{d}\theta \mathrm{d}\varphi = \delta_{ll'}\delta_{mm'}, \tag{28.3}$$

as being the eigenfunctions of angular momentum operators corresponding to different eigenvalues. The functions $\Psi_m(\varphi)$ separately are themselves orthogonal (see (27.4)), as being the eigenfunctions of the operator $\widehat{l}_z$ corresponding

to different eigenvalues m of this operator. The functions $\Theta_{lm}(\varphi)$ are not themselves eigenfunctions of any of the angular momentum operators; they are mutually orthogonal for different l, but not for different m.

The most direct method of calculating the required functions is by directly solving the problem of finding the eigenfunctions of the operator $\hat{\boldsymbol{l}}^2$ written in spherical polar coordinates (formula (26.16)). The equation $\hat{\boldsymbol{l}}^2\psi = \boldsymbol{l}^2\psi$ is:

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2} + l(l+1)\psi = 0.$$

Substituting in this equation the form (28.1) for ψ, we obtain for the function Θ_{lm} the equation

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}\Theta_{lm}}{\mathrm{d}\theta}\right) - \frac{m^2}{\sin^2\theta}\Theta_{lm} + l(l+1)\Theta_{lm} = 0. \qquad (28.4)$$

This equation is well known in the theory of spherical harmonics. It has solutions satisfying the conditions of finiteness and single-valuedness for positive integral values of $l \geqslant |m|$, in agreement with the eigenvalues of the angular momentum obtained above by the matrix method. The corresponding solutions are what are called *associated Legendre polynomials* $P_l^m(\cos\theta)$ (see §c of the Mathematical Appendices). Using the normalization condition (28.2), we find[7]

$$\Theta_{lm} = (-1)^m \mathrm{i}^l \sqrt{\frac{(2l+1)}{2}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta). \qquad (28.5)$$

Here it is supposed that $m \geqslant 0$. For negative m, we use the definition

$$\Theta_{l,-|m|} = (-1)^m\Theta_{l|m|} \qquad (28.6)$$

In other words, Θ_{lm} for $m < 0$ is given by (28.5) with $|m|$ instead of m and the factor $(-1)^m$ omitted.

Thus the angular momentum eigenfunctions are mathematically just spherical harmonic functions normalized in a particular way. For reference, the complete expression embodying the above definitions is

$$Y_{lm}(\theta,\varphi) = (-1)^{(m+|m|)/2}\mathrm{i}^l\left[\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}\right]^{1/2}P_l^{|m|}(\cos\theta)\mathrm{e}^{im\varphi}. \qquad (28.7)$$

---

[7] The choice of the phase factor is not, of course, determined by the normalization condition. The definition (28.5) used in this book is the most natural from the viewpoint of the theory of addition of angular momenta. It differs by a factor il from the one usually adopted. The advantages of this choice will be clear from the footnotes in §§60, 106 and 107.

In particular,

$$Y_{l0} = i^l \sqrt{\frac{2l + 1}{4\pi}} P_l(\cos\theta). \tag{28.8}$$

It is evident that the functions differing in the sign of m are related by

$$(-1)^{l-m} Y_{l,-m} = Y_{lm}^*. \tag{28.9}$$

For $l = 0$ (so that $m = 0$ also) the spherical harmonic function reduces to a constant. In other words, the wave functions of the states of a particle with zero angular momentum depend only on r, i.e. they have complete spherical symmetry, in agreement with the general statement in §27.

For a given m, the values of l starting from $|m|$ denumerate the successive eigenvalues of the quantity $\boldsymbol{l}^2$ in order of increasing magnitude. Hence, from the general theory of the zeros of eigenfunctions (§21), we can deduce that the function Θ_{lm} becomes zero for $l - |m|$ different values of the angle θ; in other words, it has as nodal lines $l - |m|$ "lines of latitude" on the sphere. If the complete angular functions are taken with the real factors $\cos m\varphi$ or $\sin m\varphi$ instead of[8] $e^{\pm i|m|\varphi}$, they have as further nodal lines $|m|$ "lines of longitude"; the total number of nodal lines is thus l.

Finally, we shall show how the functions Θ_{lm} may be calculated by the matrix method. This is done similarly to the calculation of the wave functions of an oscillator in §23. We start from the equation (27.8):

$$\widehat{l}_+ Y_{ll} = 0.$$

Using the expression (26.15) for the operator $\widehat{l}_+$ and substituting

$$Y_{ll} = \frac{1}{\sqrt{2\pi}} e^{il\varphi} \Theta_{ll}(\theta),$$

we obtain for Θ_{ll} the equation

$$\frac{d\Theta_{ll}}{d\theta} - l \cot\theta \cdot \Theta_{ll} = 0,$$

whence $\Theta_{ll} = \text{const} \cdot \sin^l \theta$. Determining the constant from the normalization condition, we find

$$\Theta_{ll} = (i)^l \sqrt{\frac{(2l + 1)!}{2}} \frac{1}{2^l l!} \sin^l \theta. \tag{28.10}$$

---

[8] Each such function corresponds to a state in which l_z does not have a definite value, but can have the values $\pm m$ with equal probability.

Next, using (27.12), we write

$$\widehat{l}_- Y_{l,m+1} = (l_-)_{m,m+1} Y_{lm} = \sqrt{(l-m)(l+m+1)} Y_{lm}.$$

A repeated application of this formula gives

$$\sqrt{\frac{(l-m)!}{(l+m)!}} Y_{lm} = \frac{1}{\sqrt{(2l)!}} \widehat{l}_-^{\,l-m} Y_{ll}.$$

The right-hand side of this equation is easily calculated by means of the expression (26.15) for the operator $\widehat{l}_-$. We have

$$\widehat{l}_- \left[f(\theta) e^{im\varphi} \right] = e^{i(m-1)\varphi} \sin^{1-m}\theta \frac{d}{(d\cos\theta)^{l-m}} \left(f \sin^m \theta \right)$$

A repeated application of this formula gives

$$\widehat{l}_-^{\,l-m} e^{il\varphi} \Theta_{ll} = e^{im\varphi} \sin^{-m}\theta \frac{d^{l-m}}{(d\cos\theta)^{l-m}} \left(\sin^l \theta \cdot \Theta_{ll} \right).$$

Finally, using these relations and the expression (28.10) for Θ_{ll}, we obtain the formula

$$\Theta_{lm}(\theta) = (-i)^l \sqrt{\frac{2l+1}{2} \frac{(l+m)!}{(l-m)!}} \frac{1}{2^l l! \sin^m \theta} \frac{d^{l-m}}{(d\cos\theta)^{l-m}} \sin^{2l}\theta, \qquad (28.11)$$

which is the same as (28.5).

§ 29. Matrix elements of vectors

Let us again consider a closed system of particles;[9] let f be any scalar physical quantity characterizing the system, and $\widehat{f}$ the operator corresponding to this quantity. Every scalar is invariant with respect to rotation of the coordinate system. Hence the scalar operator $\widehat{f}$ does not vary when acted on by a rotation operator, i.e. it commutes with a rotation operator. We know, however, that the operator of an infinitely small rotation is the same, apart from a constant factor, as the angular momentum operator, so that

$$\{\widehat{f}, \widehat{\boldsymbol{L}}\} = 0 \qquad (29.1)$$

---

[9]) All the results in this section are valid also for a particle in a centrally symmetric field (and in general whenever the total angular momentum of the system is conserved).

From the commutability of $\widehat{f}$ with the angular momentum operator it follows that the matrix of $\widehat{f}$ with respect to transitions between states with definite values of L and M is diagonal with respect to these suffixes. Moreover, since the specification of M defines only the orientation of the system relative to the axes of coordinates, and the value of a scalar is independent of this orientation, we can say that the matrix elements $\langle n'LM|f|nLM\rangle$ are independent of the value of M; n conventionally denotes all the quantum numbers other than L and M which define the state of the system. A formal proof of this assertion can be obtained from the commutativity of the operators $\widehat{f}$ and $\widehat{L}_+$:

$$\widehat{f}\widehat{L}_+ - \widehat{L}_+\widehat{f} = 0. \tag{29.2}$$

Let us write down the matrix element of this equation corresponding to the transition $n, L, M \to n', L, M + 1$. Taking into account the fact that the matrix of the quantity L_+ has only elements with $n, L, M \to n, L, M+1$, we obtain

$$\langle n', L, M + 1|f|n, L, M + 1\rangle\langle n, L, M + 1|L_+|n, L, M\rangle =$$
$$= \langle n', L, M + 1|L_+|n', L, M\rangle\langle n', L, M|f|n, L, M\rangle,$$

and since the matrix elements of the quantity L_+ are independent of the suffix n, we find

$$\langle n', L, M + 1|f|n, L, M + 1\rangle = \langle n', L, M|f|n, L, M\rangle, \tag{29.3}$$

whence it follows that all the quantities $\langle n', L, M|f|n, L, M\rangle$ for different M (the other suffixes being the same) are equal.

If we apply this result to the Hamiltonian itself, we obtain our previous result that the energy of the stationary states is independent of M, i.e. that the energy levels have $(2L + 1)$-fold degeneracy.

Next, let $\boldsymbol{A}$ be some vector physical quantity characterizing a closed system. When the system of coordinates is rotated (and, in particular, in an infinitely small rotation, i.e. when the angular momentum operator is applied), the components of a vector are transformed into linear functions of one another. Hence, as a result of the commutation of the operators $\widehat{L}_i$ with the operators $\widehat{A}_i$, we must again obtain components of the same vector, $\widehat{A}_i$. The exact form can be found by noticing that, in the particular case where $\boldsymbol{A}$ is the radius vector of the particle, the formulae (26.4) must be obtained. Thus we find the commutation rules

$$\{\widehat{L}_i, \widehat{A}_k\} = \mathrm{i}e_{ikl}\widehat{A}_l \tag{29.4}$$

These relations enable us to obtain several results concerning the form of the matrices of the components of the vector $\boldsymbol{A}$ (M. Born, W. Heisenberg and P. Jordan 1926). First of all, it is possible to derive selection rules which determine the transitions for which the matrix elements can be different from zero. We shall not go through the fairly lengthy calculations here, however, since it will appear later (§107) that these rules are actually a direct consequence of the general transformation properties of vector quantities and can be derived from the latter with hardly any calculation at all. Here we shall merely give the rules, without proof.

The matrix elements of all the components of a vector can be different from zero only for transitions in which the angular momentum L changes by not more than one unit:

$$L \to L, L+1 \tag{29.5}$$

There is a further selection rule which forbids transitions between any two states with $L = 0$. This rule is an obvious consequence of the complete spherical symmetry of states with angular momentum zero.

The selection rules for the angular momentum component M are different for the different components of a vector: the matrix elements can be different from zero for transitions where M changes as follows:

$$
\begin{aligned}
M \to M+1 \quad &\text{for} \quad A_+ = A_x + \mathrm{i}A_y, \\
M \to M-1 \quad &\text{for} \quad A_- = A_z - \mathrm{i}A_y, \\
M \to M \quad &\text{for} \quad A_z.
\end{aligned}
\tag{29.6}
$$

Moreover, it is possible to determine a general form for the matrix elements of a vector as functions of the number M. These important and frequently used formulae are given here, also without proof, since they are actually a particular case of more general relations derived in §107 for any tensor quantities.

The non-zero matrix elements of the quantity A_z are given by the formulae

$$\langle n'LM|A_z|nLM\rangle = \frac{M}{\sqrt{L(L+1)(2L+1)}}\langle n'L||A||nL\rangle,$$

$$\langle n'LM|A_z|n, L-1, M\rangle = \sqrt{\frac{L^2 - M^2}{L(2L-1)(2L+1)}}\langle n'L||A||n, L-1\rangle, \tag{29.7}$$

$$\langle n', L-1, M|A_z|nLM\rangle = \sqrt{\frac{L^2 - M^2}{L(2L-1)(2L+1)}}\langle n', L-1||A||nL\rangle.$$

Here the symbol

$$\langle n'L'||A||nL\rangle$$

denotes a *reduced matrix element*, a quantity independent of the quantum number M.[10]) These matrix elements are related by

$$\langle n'L'||A||nL\rangle = \langle nL||A||n'L'\rangle^*, \tag{29.8}$$

which follows directly from the fact that the operator is $\widehat{A}_z$ Hermitian.

The matrix elements of the quantities A_- and A_+ are also determined by the reduced matrix elements. The non-zero matrix elements of A_- are.

$$\langle n', L, M-1|A_-|nLM\rangle = \sqrt{\frac{(L-M+1)(L+M)}{L(L+1)(2L+1)}}\langle n'L||A||nL\rangle,$$

$$\langle n', L, M-1|A_-|n, L-1, M\rangle = \sqrt{\frac{(L-M+1)(L-M)}{L(2L-1)(2L+1)}}\langle n'L||A||n, L-1\rangle,$$

$$\langle n', L-1, M-1|A_-|n, L, M\rangle = -\sqrt{\frac{(L+M-1)(L+M)}{L(2L-1)(2L+1)}}\langle n', L-1||A||n, L\rangle \tag{29.9}$$

The matrix elements of A_+ need not be written out separately: since A_x and A_y are real we have

$$\langle n'L'M'|A_+|nLM\rangle = \langle nLM|A_-|n'L'M'\rangle^*. \tag{29.10}$$

There is a formula which expresses the matrix elements of the scalar $\boldsymbol{AB}$ in terms of the reduced matrix elements of the two vector quantities $\boldsymbol{A}$ and $\boldsymbol{B}$. The calculation is conveniently carried out by writing the operator $\widehat{\boldsymbol{A}}\widehat{\boldsymbol{B}}$ in the form

$$\widehat{\boldsymbol{A}}\widehat{\boldsymbol{B}} = \frac{1}{2}(\widehat{A}_+\widehat{B}_- + \widehat{A}_-\widehat{B}_+) + \widehat{A}_z\widehat{B}_z. \tag{29.11}$$

The matrix of $\boldsymbol{AB}$ (like that of any scalar) is diagonal with respect to L and M. A calculation by means of formulae (29.7)–(29.9) gives the result

$$\langle n'LM|\boldsymbol{AB}|nLM\rangle = \frac{1}{2L+1}\sum_{n'',L''}\langle n'L||A||n''L''\rangle\langle n''L''||B||nL\rangle, \tag{29.12}$$

where L'' takes the values $L, L \pm 1$.

---

[10]) The appearance in formulae (29.7) and (29.9) of denominators which depend on L is in accordance with the general notation used in §107. The convenience of these denominators is shown, in particular, by the simple form of equation (29.12) for the matrix elements of the scalar product of two vectors.

The symbol for the reduced matrix element is to be taken as a whole, in contrast to the matrix element symbol (see the comments following (11.17)).

For reference, we shall give the reduced matrix elements for the vector $\boldsymbol{L}$ itself. A comparison of (29.9) and (27.12) shows that

$$\langle L||L|L\rangle = \sqrt{L(L+1)(2L+1)},$$
$$\langle L-1||L||L\rangle = \langle L||L||L-1\rangle = 0. \tag{29.13}$$

A quantity that often occurs in applications is the unit vector $\boldsymbol{n}$ along the radius vector of the particle. Its reduced matrix elements can be calculated by finding, for example, the matrix elements of $n_z = \cos\theta$ for a zero angular-momentum component, $m = 0$;

$$\langle l-1, 0|n_z|l0\rangle = \int_0^\pi \Theta_{l-1,0}^* \cos\theta \cdot \Theta_{l0} \sin\theta \mathrm{d}\theta$$

with the functions Θ_{l0} given by (28.11). The evaluation of the integral[11] gives

$$\langle l-1, 0|n_z|l0\rangle = \frac{\mathrm{i}l}{\sqrt{(2l-1)(2l+1)}}.$$

The matrix elements for transitions $l \to l$ are zero (as for any polar vector of an individual particle; see (30.8) below). Comparison with (29.7) then gives

$$\langle l-1||n||l\rangle = -\langle l||n||l-1\rangle = \mathrm{i}\sqrt{l}, \quad \langle l||n||l\rangle = 0. \tag{29.14}$$

PROBLEM

Average the tensor $n_i n_k - (1/3)\delta_{ik}$ (where $\boldsymbol{n}$ is a unit vector along the radius vector of a particle) over a state where the magnitude but not the direction of the vector $\boldsymbol{1}$ is given (i.e. l_z is indeterminate).

SOLUTION. The required mean value is an operator which can be expressed in terms of the operator $\widehat{\boldsymbol{l}}$ alone. We seek it in the form

$$\overline{n_i n_k} - \frac{1}{3}\delta_{ik} = a\left[\widehat{l_i \widehat{l}_k} + \widehat{l_k \widehat{l}_i} - \frac{2}{3}\delta_{ik}l(l+1)\right];$$

this is the most general symmetrical tensor of rank two with zero trace that can be formed from the components of $\widehat{\boldsymbol{l}}$. To determine the constant a we multiply this equation on the left by $\widehat{l}_i$ and on the right by $\widehat{l}_k$ (summing over i and k). Since the vector $\boldsymbol{n}$ is perpendicular to the vector $\hbar\widehat{\boldsymbol{l}} = \widehat{\boldsymbol{r}} \times \widehat{\boldsymbol{p}}$, we have $n_i\widehat{l}_i = 0$. The product $\widehat{l}_i\widehat{l}_i\widehat{l}_k\widehat{l}_k = (\widehat{\boldsymbol{l}}^2)^2$ is replaced by its eigenvalue $l^2(l+1)^2$, and the product is transformed by means of the commutation relations (26.7) as follows:

$$\widehat{l}_i\widehat{l}_k\widehat{l}_i\widehat{l}_k = \widehat{l}_i\widehat{l}_i\widehat{l}_k\widehat{l}_k - \mathrm{i}e_{ikl}\widehat{l}_i\widehat{l}_l\widehat{l}_k = (\widehat{\boldsymbol{l}})^2 - \frac{\mathrm{i}}{2}e_{ikl}\widehat{l}_i(\widehat{l}_l\widehat{l}_k - \widehat{l}_k\widehat{l}_l) =$$

$$= (\widehat{\boldsymbol{l}})^2 + \frac{1}{2}e_{ikl}e_{lkm}\widehat{l}_i\widehat{l}_m = (\widehat{\boldsymbol{l}})^2 - \widehat{\boldsymbol{l}}^2 = l^2(l+1)^2 - l(l+1)$$

---

[11] By $l-1$ times integrating by parts with $\mathrm{d}\cos\theta$; the general formula for integrals of this type is (107.14).

(using the fact that $e_{ikl}e_{mkl} = 2\delta_{im}$). After a simple reduction we obtain the result

$$a = -\frac{1}{(2l-1)(2l+3)}.$$

§ 30. Parity of a state

Besides the parallel displacements and rotations of the coordinate system, the invariance under which represents the homogeneity and isotropy of space respectively, there is another transformation which leaves unaltered the Hamiltonian of a closed system. This is what is called the *inversion transformation*, which consists in simultaneously changing the sign of all the coordinates, i.e. a reversal of the direction of each coordinate axis; a right-handed coordinate system then becomes left-handed, and vice versa. The invariance of the Hamiltonian under this transformation expresses the symmetry of space under mirror reflections.[12] In classical mechanics, the invariance of Hamilton's function with respect to inversion does not lead to a conservation law, but the situation is different in quantum mechanics.

Let us denote by $\widehat{P}$ (for "parity") an inversion operator whose effect on a wave function $\psi(\boldsymbol{r})$ is to change the sign of the coordinates:

$$\widehat{P}\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r}) \tag{30.1}$$

It is easy to find the eigenvalues P of this operator, which are determined by the equation

$$\widehat{P}\psi(\boldsymbol{r}) = P\psi(\boldsymbol{r}) \tag{30.2}$$

To do so, we notice that a double application of the inversion operator amounts to identity: the argument of the function is unchanged. In other words, we have $\widehat{P}^2\psi = P^2\psi = \psi$, i.e. $P^2 = 1$, whence

$$P = \pm 1 \tag{30.3}$$

Thus the eigenfunctions of the inversion operator are either unchanged or change in sign when acted upon by this operator. In the first case, the wave function (and the corresponding state) is said to be *even*, and in the second it is said to be *odd*.

The invariance of the Hamiltonian under inversion (i.e. the fact that the operators $\widehat{H}$ and $\widehat{P}$ commute) thus expresses the *law of conservation of*

---

[12] Invariance under inversion exists also for the Hamiltonian of a system of particles in a centrally symmetric field with the centre at the origin.

parity: if the state of a closed system has a definite parity (i.e. if it is even, or odd) then this parity is conserved in the course of time.[13])

The angular momentum operator also is invariant under inversion, which changes the sign of the coordinates and of the operators of differentiation with respect to them; the operator (26.2) thus remains unaltered. In other words, the inversion operator commutes with the angular momentum operator, and this means that the system can have a definite parity simultaneously with definite values of the angular momentum L and its component M. All states that differ only in the value of M have the same parity; this is evident because the properties of a closed system are independent of its orientation in space, and it can be formally demonstrated from the commutation rule $\widehat{L}_+\widehat{P} - \widehat{P}\widehat{L}_+$ by the same method as in deriving (29.3) from (29.2).

There are specific *parity selection rules* for the matrix elements of various physical quantities. Let us first consider scalars. Here we must distinguish *true scalars*, which are unchanged by inversion, from *pseudoscalars*, which change sign, for instance the scalar product of an axial and a polar vector The operator of a true scalar f commutes with $\widehat{P}$; hence it follows that, if the matrix of P is diagonal, then the matrix of f is diagonal also as regards the parity suffix, i.e. the matrix elements are zero except for transitions $g \to g$ and $u \to u$ (where g and u denote even and odd states respectively). For the operator of a pseudoscalar quantity, we have $\widehat{P}\widehat{f} = -\widehat{f}\widehat{P}$; the operators $\widehat{P}$ and $\widehat{f}$ *anticommute*. The matrix element of this equation for a transition $g \to g$ is

$$P_{gg}f_{gg} = -f_{gg}P_{gg},$$

and so $f_{gg} = 0$ since $P_{gg} = 1$. Similarly we find that $f_{uu} = 0$. Thus, in the matrix of a pseudoscalar quantity, only those elements can be different from zero which correspond to transitions with change of parity. The selection rules for the matrix elements of scalars are therefore:

$$\begin{aligned}
&\text{true scalars} &&g \to g, u \to u; \\
&\text{pseudoscalars} &&g \to u, u \to g.
\end{aligned} \tag{30.4}$$

These rules can also be obtained directly from the definition of the matrix elements. Let us consider, for example, the integral $f_{ug} = \int \psi_u^* \widehat{f} \psi_g \, dq$, where the function ψ_g is even and ψ_u odd. When all the coordinates change sign, the integrand does so if f is a true scalar; on the other hand, the integral taken over all space cannot change when the variables of integration are renamed. Hence it follows that $f_{ug} = -f_{ug}$, i.e. $f_{ug} \equiv 0$.

---

[13]) To avoid misunderstanding, it should be mentioned that this refers to the non-relativistic theory. There exist interactions in Nature, falling in the realm of relativistic theory, which violate the conservation of parity.

We can similarly derive selection rules for vector quantities. Here it must be recalled that ordinary (polar) vectors change sign on inversion, while axial vectors (such as the angular momentum vector, which is the vector product of the two polar vectors $\boldsymbol{p}$ and $\boldsymbol{r}$) are unchanged by inversion. The selection rules are found to be:

$$
\begin{aligned}
&\text{polar vectors} \quad && g \to u, u \to g; \\
&\text{axial vectors} \quad && g \to g, u \to u.
\end{aligned}
\tag{30.5}
$$

Let us determine the parity of the state of a single particle with angular momentum l. The inversion transformation $(x \to -x, y \to -y, z \to -z)$ is, in spherical polar coordinates, the transformation

$$
r \to r, \quad \theta \to \pi - \theta, \quad \varphi \to \varphi + \pi.
\tag{30.6}
$$

The dependence of the wave function of the particle on the angle is given by the spherical harmonic Y_{lm}, which, apart from a constant that is here unimportant, has the form $P_l^m(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi}$. When φ is replaced by $\varphi + \pi$, the factor $\mathrm{e}^{\mathrm{i}m\varphi}$ is multiplied by $(-1)^m$, and when θ is replaced by $\pi - \theta$, $P_l^m(\cos\theta)$ becomes $P_l^m(-\cos\theta) = (-1)^{l-m}P_l^m(\cos\theta)$. Thus the whole function is multiplied by $(-1)^l$ (independent of m, in agreement with what was said above), i.e. the parity of a state with a given value of l is

$$
P = (-1)^l.
\tag{30.7}
$$

We see that all states with even l are even, and all those with odd l are odd.

A vector physical quantity relating to an individual particle can have non-zero matrix elements only for transitions with $l \to l$ or $l \pm 1$ (§29). Remembering this, and comparing formula (30.7) with what was said above regarding the change of parity in the matrix elements of vectors, we reach the result that the matrix elements of vectors relating to an individual particle are zero except for the transitions:

$$
\begin{aligned}
&\text{polar vectors} \quad && l \to l \pm 1, \\
&\text{axial vectors} \quad && l \to 1.
\end{aligned}
\tag{30.8}
$$

§ 31. Addition of angular momenta

Let us consider a system composed of two parts whose interaction is weak. If the interaction is entirely neglected, then for each part the law of conservation of angular momentum holds. The angular momentum $\boldsymbol{L}$ of the whole system can be regarded as the sum of the angular momenta $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$

of its parts. In the next approximation, when the weak interaction is taken into account, $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are not exactly conserved, but the numbers L_1 and L_2 which determine their squares remain "good" quantum numbers suitable for an approximate description of the state of the system. Regarding the angular momenta in a classical manner, we can say that in this approximation $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ rotate round the direction of L while remaining unchanged in magnitude.

For such systems the question arises regarding the "law of addition" of angular momenta: what are the possible values of L for given values of L_1 and L_2? The law of addition for the components of angular momentum is evident: since $\widehat{L}_z = \widehat{L}_{1z} + \widehat{L}_{2z}$, it follows that

$$M = M_1 + M_2 \tag{31.1}$$

There is no such simple relation for the operators of the squared angular momenta, however, and to derive their "law of addition" we reason as follows.

If we take the quantities $\boldsymbol{L}_1^2, \boldsymbol{L}_2^2, L_{1z}, L_{2z}$ as a complete set of physical quantities,[14] every state will be determined by the values of the numbers L_1, L_2, M_1, M_2. For given L_1 and L_2, the numbers M_1 and M_2 take $(2L_1 + 1)$ and $(2L_2 + 1)$ different values respectively, so that there are altogether $(2L_1 + 1)(2L_2 + 1)$ different states with the same L_1 and L_2. We denote the wave functions of the states for this representation by $\varphi_{L_1 L_2 M_1 M_2}$.

Instead of the above four quantities, we can take the four quantities $\boldsymbol{L}_1^2, \boldsymbol{L}_2^2, \boldsymbol{L}^2, L_z$ as a complete set. Then every state is characterized by the values of the numbers L_1, L_2, L, M (we denote the corresponding wave functions by $\psi_{L_1 L_2 M_1 M_2}$). For given L_1 and L_2, there must of course be $(2L_1+1)(2L_2+1)$ different states as before, i.e. for given L_1 and L_2 the pair of numbers L and M must take $(2L_1 + 1)(2L_2 + 1)$ pairs of values. These values can be determined as follows.

By adding the various possible values of M_1 and M_2, we get the corresponding values of M, as shown below:

---

[14] Together with such other quantities as form a complete set when combined with these four. These other quantities play no part in the subsequent discussion, and for brevity we shall ignore them entirely, and conventionally call the above four quantities a complete set.

M_1	M_2	M
L_1	L_2	$L_1 + L_2$
L_1	$L_2 - 1$	$L_1 + L_2 - 1$
$L_1 - 1$	L_2	$L_1 + L_2 - 1$
$L_1 - 1$	$L_2 - 1$	$L_1 + L_2 - 2$
L_1	$L_2 - 2$	$L_1 + L_2 - 2$
$L_1 - 2$	L_2	$L_1 + L_2 - 2$
...	...	...

We see that the greatest possible value of M is $M = L_1 + L_2$, corresponding to one state φ (one pair of values of M_1 and M_2). The greatest possible value of M in the states ψ, and hence the greatest possible value of L also, is therefore $L_1 + L_2$. Next, there are two states φ with $M = L_1 + L_2 - 1$. Consequently, there must also be two states ψ with this value of M; one of them is the state with $L = L_1 + L_2$ (and $M = L - 1$), and the other is that with $L = L_1 + L_2 - 1$ (and $M = L$). For the value $M = L_1 + L_2 - 2$ there are three different states φ. This means that, besides the values $L = L_1 + L_2, L = L_1 + L_2 - 1$, the value $L = L_1 + L_2 - 2$ can occur.

The argument can be continued in this way so long as a decrease of M by 1 increases by 1 the number of states with a given M. It is easily seen that this is so until M reaches the value $|L_1 - L_2|$. When M decreases further, the number of states no longer increases, remaining equal to $2L_2 + 1$ (if $L_2 \leqslant L_1$). Thus $|L_1 - L_2|$ is the least possible value of L, and we arrive at the result that, for given L_1 and L_2, the number L can take the values

$$L = L_1 + L_2, L_1 + L_2 - 1, \ldots, |L_1 - L_2|, \qquad (31.2)$$

that is $2L_2 + 1$ different values altogether (supposing that $L_2 \leqslant L_1$). It is easy to verify that we do in fact obtain $(2L_1 + 1)(2L_2 + 1)$ different values of the pair of numbers M, L. Here it is important to note that, if we ignore the $2L + 1$ values of M for a given $\widehat{L}$, then only one state will correspond to each of the possible values (31.2) of L.

This result can be illustrated by means of what is called the *vector model*. If we take two vectors $\boldsymbol{L}_1, \boldsymbol{L}_2$ of lengths L_1 and L_2, then the values of L are represented by the integral lengths of the vectors $\boldsymbol{L}$ which are obtained by vector addition of $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$; the greatest value of L is $L_1 + L_2$, which is obtained when $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are parallel, and the least value is $|L_1 - L_2|$, when $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are antiparallel.

In states with definite values of the angular momenta L_1, L_2 and of the total angular momentum L, the scalar products $\boldsymbol{L}_1\boldsymbol{L}_2$, $\boldsymbol{L}\boldsymbol{L}_1$ and $\boldsymbol{L}\boldsymbol{L}_2$ also

have definite values. These values are easily found. To calculate $\boldsymbol{L}_1\boldsymbol{L}_2$, we write $\widehat{\boldsymbol{L}} = \widehat{\boldsymbol{L}}_1 + \widehat{\boldsymbol{L}}_2$ or, squaring and transposing,

$$2\widehat{\boldsymbol{L}}_1\widehat{\boldsymbol{L}}_2 = \widehat{\boldsymbol{L}}^2 - \widehat{\boldsymbol{L}}_1^2 - \widehat{\boldsymbol{L}}_2^2.$$

Replacing the operators on the right-hand side of this equation by their eigenvalues, we obtain the eigenvalue of the operator on the left-hand side:

$$\boldsymbol{L}_1\boldsymbol{L}_2 = \frac{1}{2}\{L(L+1) - L_1(L_1+1) - L_2(L_2+1)\}. \tag{31.3}$$

Similarly we find

$$\boldsymbol{L}\boldsymbol{L}_1 = \frac{1}{2}\{L(L+1) + L_1(L_1+1) - L_2(L_2+1)\} \tag{31.4}$$

Let us now determine the "addition rule for parities". The wave function Ψ of a system consisting of two independent parts is the product of the wave functions ψ_1 and Ψ_2 of these parts. Hence it is clear that if the latter are of the same parity (i.e. both change sign, or both do not change sign, when the sign of all the coordinates is reversed), then the wave function of the whole system is even. On the other hand, if Ψ_1 and Ψ_2 are of opposite parity, then the function Ψ is odd. These statements may be written

$$P = P_1 P_2 \tag{31.5}$$

where P is the parity of the whole system and P_1, P_2 those of its parts. This rule can, of course, be generalized at once to the case of a system composed of any number of non-interacting parts.

In particular, if we are concerned with a system of particles in a centrally symmetric field (the mutual interaction of the particles being supposed weak), then the parity of the state of the whole system is given by

$$P = (-1)^{l_1+l_2+\cdots} \tag{31.6}$$

see (30.7). We emphasize that the exponent here contains the algebraic sum of the angular momenta l_i, and this is not in general the same as their "vector sum", i.e. the angular momentum L of the system.

If a closed system disintegrates (under the action of internal forces), the total angular momentum and parity must be conserved. This circumstance may render it impossible for a system to disintegrate, even if this is energetically possible.

For instance, let us consider an atom in an even state with angular momentum $L = 0$, which is able, so far as energy considerations go, to disintegrate into a free electron and an ion in an odd state with the same angular

momentum $L = 0$. It is easy to see that in fact no such disintegration can occur (it is, as we say, *forbidden*). For, by virtue of the law of conservation of angular momentum, the free electron would also have to have zero angular momentum, and therefore be in an even state ($P = (1)^0 = +1$); the state of the system ion+electron would then be odd, however, whereas the original state of the atom was even.

CHAPTER V

MOTION IN A CENTRALLY SYMMETRIC FIELD

121

CHAPTER VI

PERTURBATION THEORY

§ 38. Perturbations independent of time

THE exact solution of Schrödinger's equation can be found only in a comparatively small number of the simplest cases. The majority of problems in quantum mechanics lead to equations which are too complex to be solved exactly. Often, however, quantities of different orders of magnitude appear in the conditions of the problem; among them there may be small quantities such that, when they are neglected, the problem is so much simplified that its exact solution becomes possible. In such cases, the first step in solving the physical problem concerned is to solve exactly the simplified problem, and the second step is to calculate approximately the errors due to the small terms that have been neglected in the simplified problem. There is a general method, of calculating these errors; it is called *perturbation theory*.

Let us suppose that the Hamiltonian of a given physical system is of the form

$$\widehat{H} = \widehat{H}_0 + \widehat{V}$$

where $\widehat{V}$ is a small correction (or *perturbation*) to the *unperturbed* operator $\widehat{H}_0$ In §§40, 39 we shall consider perturbations $\widehat{V}$ which do not depend explicitly on time (the same is assumed regarding $\widehat{H}_0$ also). The conditions which are necessary for it to be permissible to regard the operator $\widehat{V}$ as "small" compared with the operator $\widehat{H}$ will be derived below.

The problem of perturbation theory for a discrete spectrum can be formulated as follows. It is assumed that the eigenfunctions $\psi_n^{(0)}$ and eigenvalues $E_n^{(0)}$ of the discrete spectrum of the unperturbed operator $\widehat{H}_0$ are known, i.e. the exact solutions of the equation

$$\widehat{H}_0 \psi^{(0)} = E^{(0)} \psi^{(0)} \tag{38.1}$$

are known. It is desired to find approximate solutions of the equation

$$\widehat{H}\psi = (\widehat{H}_0 + \widehat{V})\psi = E\psi, \tag{38.2}$$

i.e. approximate expressions for the eigenfunctions ψ_n and eigenvalues E_n of the perturbed operator $\widehat{H}$.

In this section we shall assume that no eigenvalue of the operator $\widehat{H}_0$ is degenerate. Moreover, to simplify our results, we shall at first suppose that there is only a discrete spectrum of energy levels.

The calculations are conveniently performed in matrix form throughout.

To do this, we expand the required function ψ in terms of the functions $\psi_n^{(0)}$:

$$\psi = \sum_m c_m \psi_m^{(0)}, \tag{38.3}$$

Substituting this expansion in (38.2) we obtain

$$\sum_m c_m (E_m^{(0)} + \widehat{V}) \psi_m^{(0)} = \sum_m c_m E \psi_m^{(0)};$$

multiplying both sides of this equation by $\psi_k^{(0)*}$ and integrating, we find

$$(E - E_k^{(0)}) c_k = \sum_m V_{km} c_m. \tag{38.4}$$

Here we have introduced the matrix V_{km} of the perturbation operator $\widehat{V}$, defined with respect to the unperturbed functions $\psi_m^{(0)}$:

$$V_{km} = \int \psi_k^{(0)*} \widehat{V} \psi_m^{(0)} \mathrm{d}q. \tag{38.5}$$

We shall seek the values of the coefficients c_m and the energy E in the form of series

$$E = E^{(0)} + E^{(1)} + \ldots, \quad c_m = c_m^{(0)} + c_m^{(1)} + c_m^{(2)} + \ldots,$$

where the quantities $E^{(1)}$ and $c_m^{(1)}$ are of the same order of smallness as the perturbation $\widehat{V}$, the quantities $E^{(2)}$ and $c_m^{(2)}$ are of the second order of smallness, and so on.

Let us determine the corrections to the nth eigenvalue and eigenfunction, putting accordingly $c_n^{(0)} = 1$, $c_n^{(0)} = 0$ for $m \neq n$. To find the first approximation, we substitute in equation (38.4) $E = E_n^{(0)} + E_n^{(1)}$, $c_k = c_k^{(0)} + c_k^{(1)}$, and retain only terms of the first order. The equation with $k = n$ gives

$$E_n^{(1)} = V_{nn} = \int \psi_n^{(0)*} \widehat{V} \psi_n^{(0)} \mathrm{d}q. \tag{38.6}$$

Thus the first-order correction to the eigenvalue $E_n^{(0)}$ is equal to the mean value of the perturbation in the state $\psi_n^{(0)}$.

The equation (38.4) with $k \neq n$ gives

$$c_k^{(1)} = \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}, k \neq n, \tag{38.7}$$

while $c_n^{(1)}$ remains arbitrary; it must be chosen so that the function $\psi_n = \psi_n^{(0)} + \psi_n^{(1)}$ is normalized up to and including terms of the first order. For this we must put $c_n^{(1)} = 0$. For the functions

$$\psi_n^{(1)} = {\sum_m}' \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} \tag{38.8}$$

(the prime means that the term with $m = n$ is omitted from the sum) are orthogonal to ψ_n^0, and hence the integral of $|\psi_n^{(0)} + \psi_n^{(1)}|^2$ differs from unity only by a quantity of the second order of smallness.

Formula (38.8) determines the correction to the wave functions in the first approximation. Incidentally, we see from this formula the condition for the applicability of the above method. This condition is that the inequality

$$|V_{mn}| \ll |E_n^{(0)} - E_m^{(0)}| \tag{38.9}$$

must hold, i.e. the matrix elements of the perturbation must be small compared with the corresponding differences between the unperturbed energy levels.

Next, let us determine the correction to the eigenvalue $E_n^{(0)}$ in the second approximation. To do this, we substitute in (38.4) $E = E_n^{(0)} + E_n^{(1)} + E_n^{(2)}$, $c_k = c_k^{(0)} + c_k^{(1)} + c_k^{(2)}$, and examine the terms of the second order of smallness. The equation with $k = n$ gives

$$E_n^{(2)} c_n^0 = {\sum_m}' V_{nm} c_m^{(1)},$$

whence

$$E_n^{(2)} = {\sum_m}' \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} \tag{38.10}$$

(we have substituted $c_m^{(1)}$ from (38.7), and used the fact that, since the operator $\widehat{V}$ is Hermitian, $V_{mn} = V_{nm}^*$).

We notice that the correction in the second approximation to the energy of the normal state is always negative; for, since $E_n^{(0)}$ then corresponds to the lowest value of the energy, all the terms in the sum (38.10) are negative.

The further approximations can be calculated in a similar manner.

The results obtained can be generalized at once to the case where the operator $\widehat{H}_0$ has also a continuous spectrum (but the perturbation is applied, as before, to a state of the discrete spectrum). To do so, we need only add to the sums over the discrete spectrum the corresponding integrals over the continuous spectrum. We shall distinguish the various states of the continuous spectrum by the suffix ν, which takes a continuous range of values; by ν we conventionally understand an assembly of values of quantities sufficient for a complete description of the state (if the states of the continuous spectrum are degenerate, which is almost always the case, the value of the energy alone does not suffice to determine the state).[1]) Then, for instance, we must write instead of (38.8)

$$\psi_n^{(1)} = {\sum_m}' \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} + \int \frac{V_{\nu n}}{E_n^{(0)} - E_\nu} \psi_\nu^{(0)} \mathrm{d}\nu \qquad (38.11)$$

and similarly for the other formulae.

It is useful to note also the formula for the perturbed value of the matrix element of a physical quantity f, calculated as far as terms of the first order by using the functions $\psi_n = \psi_n^{(0)} + \psi_n^{(1)}$, with $\psi_n^{(1)}$ given by (38.8). The following expression is easily obtained:

$$f_{nm} = f_{nm}^{(0)} + {\sum_k}' \frac{V_{nk} f_{km}^{(0)}}{E_n^{(0)} - E_k^{(0)}} + {\sum_k}' \frac{V_{km} f_{nk}^{(0)}}{E_m^{(0)} - E_k^{(0)}} \qquad (38.12)$$

In the first sum $k \neq n$, while in the second $k \neq m$.

PROBLEMS

1. Determine the correction $\psi_n^{(2)}$ in the second approximation to the eigenfunctions.

SOLUTION. The coefficients $c_k^{(2)} (k \neq n)$ are calculated from equations (38.4) with $k \neq n$, written out up to terms of the second order, and the coefficient $c_n^{(2)}$ is chosen so that the function $\psi_n = \psi_n^{(0)} + \psi_n^{(1)} + \psi_n^{(2)}$ is normalized up to terms of the second order. As a result we find

$$\psi_n^{(2)} = {\sum_m}'{\sum_k}' \frac{V_{mk} V_{kn}}{\hbar^2 \omega_{nk} \omega_{nm}} \psi_m^{(0)} - {\sum_m}' \frac{V_{nn} V_{mn}}{\hbar^2 \omega_{nm}^2} \psi_m^{(0)} - \frac{\psi_n^{(0)}}{2} {\sum_m}' \frac{|V_{mn}|^2}{\hbar^2 \omega_{nm}^2},$$

where we have introduced the frequencies

$$\omega_{nm} = \frac{1}{\hbar}(E_n^{(0)} - E_m^{(0)}).$$

---

[1]) Here the wave functions $\psi_\nu^{(0)}$ must be normalized by delta functions of the quantities ν.

2. Determine the correction in the third approximation to the eigenvalues of the energy

SOLUTION. Writing out the terms of the third order of smallness in equation (38.4) with $k = n$, we obtain

$$E_n^{(3)} = \sum_k {}' \sum_m {}' \frac{V_{nm} V_{mk} V_{kn}}{\hbar^2 \omega_{mn} \omega_{kn}} - V_{nn} \sum_m {}' \frac{|V_{nm}|^2}{\hbar^2 \omega_{mn}^2}.$$

3. Determine the energy levels of an anharmonic linear oscillator whose Hamiltonian is

$$\widehat{H} = \frac{\widehat{p}^2}{2m} + \frac{m\omega^2 x^2}{2} + \alpha x^3 + \beta x^4.$$

SOLUTION. The matrix elements of x^3 and $^x4$ can be obtained directly according to the rule of matrix multiplication, using the expression (23.4) for the matrix elements of x. We find for the matrix elements of x^3 that are not zero

$$\left(x^3\right)_{n-3,n} = \left(x^3\right)_{n,n-3} = \left(\frac{\hbar}{m\omega}\right)^{3/2} \sqrt{\frac{n(n-1)(n-2)}{8}},$$

$$\left(x^3\right)_{n-1,n} = \left(x^3\right)_{n,n-1} = \left(\frac{\hbar}{m\omega}\right)^{3/2} \sqrt{\frac{9n^3}{8}}.$$

The diagonal elements in this matrix vanish, so that the correction in the first approximation due to the term αx^3 in the Hamiltonian (regarded as a perturbation of the harmonic oscillator) is zero. The correction in the second approximation due to this term is of the same order as that in the first approximation due to the term βx^4. The diagonal matrix elements of x^4 are

$$\left(x^4\right)_{n,n} = \left(\frac{\hbar}{m\omega}\right)^2 \cdot \frac{3}{4}(2n^2 + 2n + 1).$$

Using the general formulae (38.6) and (38.10), we find the following approximate expression for the energy levels of the anharmonic oscillator:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) - \frac{15}{4}\frac{\alpha^2}{\hbar\omega}\left(\frac{\hbar}{m\omega}\right)^3\left(n^2 + n + \frac{11}{30}\right) + \frac{3}{2}\beta\left(\frac{\hbar}{m\omega}\right)^2\left(n^2 + n + \frac{1}{2}\right).$$

4. A spherical potential well with infinitely high walls is subjected to a small deformation (without change of volume) which gives it the form of a slightly prolate or oblate spheroid with semi-axes $a = b$ and c. Find the splitting of the energy levels of a particle in the deformed well (A. B. Migdal 1959).

SOLUTION. The equation of the well boundary is

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1$$

and by the change of variables $x \to ax/R, y \to ay/R, z \to cz/R$ it is converted into $x^2 + y^2 + z^2 = R^2$ the equation of a sphere with radius R. The same change

of variables converts the Hamiltonian of the particle, $\widehat{H} = \widehat{\boldsymbol{p}}^2/2M = -\hbar^2\Delta/2M$ (where M is the mass of the particle and the energy is measured from the bottom of the well) into $\widehat{H} = \widehat{H}_0 + \widehat{V}$, where

$$\widehat{H}_0 = -\frac{\hbar^2}{2M}\Delta, \quad \widehat{V} = -\frac{\hbar^2}{2M}\left[\left(\frac{R^2}{a^2} - 1\right)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \left(\frac{R^2}{c^2} - 1\right)\frac{\partial^2}{\partial z^2}\right].$$

Thus the problem of motion in an ellipsoidal well reduces to that of motion in a spherical well. If the ellipsoid is almost a sphere of radius $R = (a^2c)^{1/3}$, $\widehat{V}$ may be regarded as a small perturbation. If the ellipsoidality $\beta(|\beta| \ll 1)$ is defined by

$$\alpha \approx R\left(1 - \frac{\beta}{3}\right), \quad c \approx R\left(1 + \frac{2\beta}{3}\right),$$

the perturbation operator may be written

$$\widehat{V} = \frac{\beta}{3M}(\widehat{\boldsymbol{p}}^2 - 3\widehat{p}_z^2).$$

In the first order of perturbation theory, the change in the energy levels of the particle from their values in the spherical well is

$$\Delta E_{nlm} = E_{nlm} - E_{nl}^{(0)} = \langle nlm|V|nlm\rangle$$

where l and m are the angular momentum of the particle and its component along the axis of the spheroid; n numbers the levels in the spherical well for a given l, which are independent of m. Since $\boldsymbol{p}^2 - 3p_z^2$ is the zz-component of an irreducible tensor, $\delta_{ik}\boldsymbol{p}^2 - 3p_ip_k$, with zero trace, we find from (107.2) and (107.6) that the matrix element $\langle nlm|V|nln\rangle$ is proportional to

$$(-1)^m \begin{pmatrix} l & 2 & l \\ -m & 0 & m \end{pmatrix},$$

and therefore

$$\langle nlm|V|nlm\rangle = \left(1 - \frac{3m^2}{l(l+1)}\right)\langle nl0|V|nl0\rangle$$

A table of $3j$-symbols is given in §106.

Next,

$$\langle nl0|V|nl0\rangle = \frac{2}{3}\beta E_{nl}^{(0)} + \beta\frac{\hbar^2}{M}\left\langle nl0\left|\frac{\partial^2}{\partial z^2}\right|nl0\right\rangle =$$

$$= \frac{2}{3}\beta E_{nl}^{(0)} - \frac{\beta\hbar^2}{M}\int\left|\frac{\partial\psi_{nl0}}{\partial z}\right|^2 r^2\mathrm{d}r\mathrm{d}o$$

in the first term we have used Schrödinger's equation $\widehat{H}_0\psi_{nlm} = E_{nl}^{(0)}\psi_{nlm}$ for a spherical well, and in the second term integrated by parts. With Y_{l0} in the form

(28.11), we find the derivative of $\psi_{nl0} = R_{nl}(r)Y_{l0}(\theta, \varphi)$ to be

$$\frac{\partial}{\partial z}\psi_{nl0} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\psi_{nl0} =$$

$$= -\frac{i(l+1)}{[4(l+1)^2-1]^{1/2}}\left(R'_{nl} - \frac{l}{r}R_{nl}\right)Y_{l+1,0}+$$

$$+ \frac{il}{[4l^2-1]^{1/2}}\left(R'_{nl} + \frac{l+1}{r}R_{nl}\right)Y_{l-1,0}.$$

The radial integrals are calculated by means of the formulae

$$\int_0^\infty R_{nl}R'_{nl}r\mathrm{d}r = -\frac{1}{2}\int_0^\infty R^2_{nl}\mathrm{d}r.$$

$$\int_0^\infty R'^2_{nl}r^2\mathrm{d}r = \frac{2M}{\hbar^2}E^{(0)}_{nl} - l(l+1)\int_0^\infty R^2_{nl}\mathrm{d}r,$$

which are derived by integrating by parts and using the radial Schrödinger's equation (**??**)

$$R''_{nl} + \frac{2}{r}R''_{nl} - \frac{l(l+1)}{r^2}R_{nl} = -\frac{2M}{\hbar^2}E^{(0)}_{nl}.$$

The terms containing integrals of R^2_{nl} cancel, and the final result is

$$\Delta E_{nlm} = 4\beta\frac{l(l+1)}{(2l-1)(2l+3)}\left[\frac{m^2}{l(l+1)} - \frac{1}{3}\right]E^{(0)}_{nl}.$$

Note that

$$\frac{1}{2l+1}\sum_{m=-l}^{l}E_{nlm} = E^{(0)}_{nl},$$

i.e. the "centre of gravity" of the multiplet is not shifted.

§ 39. The secular equation

Let us now turn to the case where the unperturbed operator $\widehat{H}_0$ has degenerate eigenvalues. We denote by $\psi_n^{(0)}, \psi_{n'}^{(0)}, \ldots$ the eigenfunctions belonging to the same eigenvalue $E_n^{(0)}$ of the energy. The choice of these functions is, as we know, not unique; instead of them we can choose any s (where s is the degree of degeneracy of the level $E_n^{(0)}$) independent linear combinations of these functions. The choice ceases to be arbitrary, however, if we subject the wave functions to the requirement that the change in them under the action of the small applied perturbation should be small.

At present we shall understand by $\psi_n^{(0)}, \psi_{n'}^{(0)}, \ldots$ some arbitrarily selected unperturbed eigenfunctions. The correct functions in the zeroth approximation are linear combinations of the form

$$c_n^{(0)} \psi_n^{(0)} + c_{n'}^{(0)} \psi_{n'}^{(0)} + \ldots$$

The coefficients in these combinations are determined, together with the corrections in the first approximation to the eigenvalues, as follows.

We write out equations (38.4) with $k = n, n', \ldots$, and substitute in them, in the first approximation, $E = E_n^{(0)} + E^{(1)}$; for the quantities c_k it suffices to take the zero-order values $c_n = c_n^{(0)}, c_{n'} = c_{n'}^{(0)}, \ldots ; c_m = 0$ for $m \neq n, n', \ldots$. We then obtain

$$E^{(1)} c_n^{(0)} = \sum_{n'} V_{nn'} c_{n'}^{(0)}$$

or

$$\sum_{n'} \left(V_{nn'} - E^{(1)} \delta_{nn'} \right) c_{n'}^{(0)} = 0, \tag{39.1}$$

where n, n' take all values denumerating states belonging to the given unperturbed eigenvalue $E_n^{(0)}$. This system of homogeneous linear equations for the quantities $c_n^{(0)}$ has solutions which are not all zero if the determinant of the coefficients of the unknowns vanishes. Thus we obtain the equation

$$\left| V_{nn'} - E^{(1)} \delta_{nn'} \right| = 0. \tag{39.2}$$

This equation is of the sth degree in $E^{(1)}$ and has, in general, s different real roots. These roots are the required corrections to the eigenvalues in the first approximation. Equation (39.2) is called the *secular equation.*[2]) We notice that the sum of its roots is equal to the sum of the diagonal matrix elements $V_{nn}, V_{n'n'}, \ldots$ (this being the coefficient of $[E^{(1)}]^{s-1}$ in the equation).

Substituting in turn the roots of equation (39.2) in the system (39.1) and solving, we find the coefficients $c_n^{(0)}$ and so determine the eigenfunctions in the zeroth approximation.

As a result of the perturbation, an originally degenerate energy level ceases in general to be degenerate (the roots of equation (39.2) are in general distinct); the perturbation removes the degeneracy, as we say. The removal of the degeneracy may be either total or partial (in the latter case, after the perturbation has been applied, there remains a degeneracy of degree less than the original one).

It may happen that for some reason all the matrix elements are particularly small (or even zero) for transitions within a group of mutually degenerate states $n, n', \ldots$. It may then be useful to take into account not only

---

[2]) The name is taken from celestial mechanics.

in the first order the matrix elements $V_{nn'}$. but also in the higher orders the matrix elements $V_{nm}(m \neq n, n', \dots)$ for transitions to states with a different energy. Let us do this for the matrix elements Vmn in the second order.

In equation (38.4) with $k = n$ we put on the left $E = E_n^{(0)} + E_n^{(1)}$ (retaining the notation $E^{(1)}$ for the correction to the energy in the approximation considered), and replace c_n by $c_n^{(0)}$. Since $c_m^{(0)} = 0$ for all $m \neq n, n'$ we have

$$E^{(1)}c_n^{(0)} = \sum_m V_{nm}c_m^{(1)} + \sum_{n'} V_{nn'}c_{n'}^{(0)}. \tag{39.3}$$

The equations (38.4) with $k = m \neq n, n', \dots$ give as far as the first-order terms

$$(E_n^{(0)} - E_m^{(0)})c_m^{(1)} = \sum_{n'}^{(1)} = \sum_{n'} V_{mn'}c_{n'}^{(0)},$$

whence

$$c_m^{(1)} = \sum_{n'} \frac{V_{mn'}}{E_n^{(0)} - E_m^{(0)}}c_{n'}^{(0)}.$$

Substitution in (39.3) gives

$$E^{(1)}c_n^{(0)} = \sum_{n'} c_{n'}^{(0)} \left(V_{nn'} + \sum_m \frac{V_{nm}V_{mn'}}{E_n^{(0)} - E_m^{(0)}} \right).$$

These equations replace (39.1); the condition for them to be compatible again leads to the secular equation, which differs from (39.2) by the change

$$V_{nn'} \to V_{nn'} + \sum_m \frac{V_{nm}V_{mn'}}{E_n^{(0)} - E_m^{(0)}} \tag{39.4}$$

PROBLEMS

1. Determine the corrections to the eigenvalue in the first approximation and the correct functions in the zeroth approximation, for a doubly degenerate level.

SOLUTION. Equation (39.2) here has the form

$$\begin{vmatrix} V_{11} - E^{(1)} & V_{21} \\ V_{12} & V_{22} - E^{(1)} \end{vmatrix} = 0$$

(the suffixes 1 and 2 correspond to two arbitrarily chosen unperturbed eigenfunctions $\psi_1^{(0)}$ and $\psi_1^{(0)}$ of the degenerate level in question). Solving, we find

$$E^{(}1) = \frac{1}{2}\left[V_{11} + V_{22} \pm \hbar\omega^{(1)} \right], \tag{1}$$

with the notation

$$\hbar\omega^{(1)} = \sqrt{(V_{11} - V_{22})^2 + 4|V_{12}|^2}$$

for the difference between the two values of the correction $E^{(1)}$ Solving also equations (39.1) with these values of $E^{(1)}$, we obtain for the coefficients in the correct normalized function in the zeroth approximation, $\psi^{(0)} = c_1^{(0)}\psi_1^{(0)} + c_2^{(0)}\psi_2^{(0)}$, the values

$$c_1^{(0)} = \left\{ \frac{V_{12}}{2|V_{12}|} \left[1 \pm \frac{V_{11} - V_{22}}{\hbar\omega^{(1)}} \right] \right\}^{1/2},$$
$$c_2^{(0)} = \pm \left\{ \frac{V_{21}}{2|V_{12}|} \left[1 \mp \frac{V_{11} - V_{22}}{\hbar\omega^{(1)}} \right] \right\}^{1/2}. \tag{2}$$

2. Derive the formulae for the correction to the eigenfunctions in the first approximation and to the eigenvalues in the second approximation.

SOLUTION. We shall suppose that the correct functions in the zeroth approximation are chosen as the functions $\psi_n^{(0)}$. The matrix $V_{nn'}$, defined with respect to these is clearly diagonal with respect to the suffixes n, n' (belonging to the same group of functions of a degenerate level), and the diagonal elements $V_{nn}, V_{n'n'}$, are equal to the corresponding corrections $E_n^{(1)}, E_{n'}^{(1)}, \dots$ in the first approximation.

Let us consider a perturbation of the eigenfunction $\psi_n^{(0)}$, so that in the zeroth approximation $E = E_n^{(0)}, c_n^{(0)} = 1, c_m^{(0)} = 0$ for $m \neq n$. In the first approximation $E = E_n^{(0)} + V_{nn}, c_n = 1 + c_n^{(1)}, c_m = c_m^{(1)}$. We write out from the system (38.4) the equation with $k \neq n, n', \dots$, retaining in it terms of the first order:

$$(E_n^{(0)} - E_k^{(0)})c_k^{(1)} = V_{kn}c_n^{(0)} = V_{kn},$$

whence

$$c_k^{(1)} = \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} \quad \text{for } k \neq n, n', \dots \tag{1}$$

Next we write out the equation with $k = n'$, retaining in it terms of the second order:

$$E_n^{(1)}c_{n'}^{(1)} = V_{n'n'}c_{n'}^{(1)} + {\sum_m}' V_{n'm}c_m^{(1)}$$

(the terms with $m = n, n', \dots$ are omitted in the sum over m). Substituting $E_n^{(1)} = V_{nn}$ and the expression (1) for $c_m^{(1)}$, we obtain for $n' \neq n$

$$c_{n'}^{(1)} = \frac{1}{(V_{nn} - V_{n'n'})} {\sum_m}' \frac{V_{n'm}V_{mn}}{E_n^{(0)} - E_m^{(0)}} \tag{2}$$

(In this approximation the coefficient $c_n^{(1)}$ is zero.) Formulae (1) and (2) determine

the correction $\psi_n^{(1)} = \sum_m c_m^{(1)} \psi_m^{(0)}$ to the eigenfunctions in the first approximation.[3])

Finally, writing out the second-order terms in equation (38.4) with $k = n$, we obtain for the second-order corrections to the energy the formula

$$E_n^{(2)} = \sum_m{}' \frac{V_{nm} V_{mn}}{E_n^{(0)} - E_m^{(0)}} \tag{3}$$

which is formally identical with (38.10).

3. At the initial instant $t = 0$, a system is in a state $\psi_1^{(0)}$ which belongs to a doubly degenerate level. Determine the probability that, at a subsequent instant t, the system will be in the state $\psi_2^{(0)}$ with the same energy; the transition occurs under the action of a constant perturbation.

SOLUTION. We form the correct functions in the zeroth approximation,

$$\psi = c_1 \psi_1 + c_2 \psi_2, \quad \psi' = c_1' \psi_1 + c_2' \psi_2,$$

where $c_1, c_2; c_1', c_2'$ are two pairs of coefficients determined by formulae (1) of Problem 1 (for brevity, we omit the index (0) on all quantities).

Conversely,

$$\psi_1 = \frac{c_2' \psi - c_2 \psi'}{c_1 c_2' - c_1' c_2}.$$

The functions ψ and ψ' belong to states with perturbed energies $E + E^{(1)}$ and $E + E^{(1)'}$, where $E^{(1)}$ and $E^{(1)'}$ are the two values of the correction (1) in Problem 1. On introducing the time factors we pass to the time-dependent wave functions:

$$\psi_1 = \frac{\exp(-(\mathrm{i}/\hbar)Et)}{c_1 c_2' - c_1' c_2} \left[c_2' \exp\left(-\frac{\mathrm{i}}{\hbar} E^{(1)} t\right) - c_2 \psi' \exp\left(-\frac{\mathrm{i}}{\hbar} E^{(1)'} t\right) \right]$$

(at time $t = 0, \Psi_1 = \psi_1$). Finally, again expressing ψ, ψ' in terms of ψ_1, ψ_2, we obtain Ψ_1, as a linear combination of ψ_1, and ψ_2, with coefficients depending on time. The squared modulus of the coefficient of ψ_2 determines the required transition probability w_{21}. Calculation with (1) and (2) from Problem 1 gives

$$w_{21} = 2 \frac{|V_{12}|^2}{(\hbar \omega^{(1)})^2} [1 - \cos(\omega^{(1)} t)].$$

We see that the probability varies periodically with time, with frequency $\omega^{(1)}$.

---

[3]) Note that the condition for the quantities (1) and (2) to be small (and therefore the condition for this method of perturbation theory to be applicable) again requires the conditions (38.9) to be satisfied only for transitions between states belonging to different energy levels. Transitions between states belonging to the same degenerate level are taken into account exactly (in a certain sense) by the secular equation.

For times t which are small compared with the period in question, the expression in the braces, and therefore w_{21}, is proportional to t^2:

$$w_{21} = \frac{1}{\hbar^2}|V_{21}|^2 t^2;$$

This formula can be very simply obtained by the method given in the next section (using equation (40.4)).

§ 40. Perturbations depending on time

Let us now go on to study perturbations depending explicitly on time. We cannot speak in this case of corrections to the eigenvalues, since, when the Hamiltonian is time-dependent (as will be the perturbed operator $\widehat{H} = \widehat{H}_0 + \widehat{V}(t)$), the energy is not conserved, so that there are no stationary states. The problem here consists in approximately calculating the wave functions from those of the stationary states of the unperturbed system.

To do this, we shall apply a method analogous to the well-known method of varying the constants to solve linear differential equations (P. A. M. Dirac 1926). Let $\Psi_k^{(0)}$ be the wave functions (including the time factor) of the stationary states of the unperturbed system. Then an arbitrary solution of the unperturbed wave equation can be written in the form of a sum $\Psi = \sum a_k \Psi_k^{(0)}$. We shall now seek the solution of the perturbed equation

$$i\hbar\frac{\partial \Psi}{\partial t} = (\widehat{H}_0 + \widehat{V})\Psi \tag{40.1}$$

in the form of a sum

$$\Psi = \sum_k a_k(t)\Psi_k^{(0)}, \tag{40.2}$$

where the expansion coefficients are functions of time. Substituting (40.2) in (40.1), and recalling that the functions $\Psi_k^{(0)}$ satisfy the equation

$$i\hbar\frac{\partial \Psi_k^{(0)}}{\partial t} = \widehat{H}_0\Psi_k^{(0)},$$

we obtain

$$i\hbar\sum_k \Psi_k^{(0)}\frac{\mathrm{d}a_k}{\mathrm{d}t} = \sum_k a_k\widehat{V}\Psi_k^{(0)}.$$

Multiplying both sides of this equation on the left by $\Psi_m^{(0)*}$ and integrating, we have

$$i\hbar\frac{\mathrm{d}a_m}{\mathrm{d}t} = \sum_k V_{mk}(t)a_k, \tag{40.3}$$

where

$$V_{mk}(t) = \int \Psi_m^{(0)*} \widehat{V} \Psi_k^{(0)} \mathrm{d}q = V_{mk} \mathrm{e}^{\mathrm{i}\omega_{mk}t}, \quad \omega_{mk} = \frac{E_m^{(0)} - E_k^{(0)}}{\hbar}$$

are the matrix elements of the perturbation, including the time factor (and it must be borne in mind that, when V depends explicitly on time, the quantities V_{mk} also are functions of time).

As the unperturbed wave function we take the wave function of the nth stationary state, for which the corresponding values of the coefficients in (40.2) are $a_n^{(0)} = 1, a_k^{(0)} = 0$ for $k \neq n$. To find the first approximation, we seek a_k in the form $a_k = a_k^{(0)} + a_k^{(1)}$ substituting $a_k = a_k^{(0)}$ on the right-hand side of equation (40.3), which already contains the small quantities V_{mk}. This gives

$$\mathrm{i}\hbar \frac{\mathrm{d}a_k^{(1)}}{\mathrm{d}t} = V_{kn}(t) \tag{40.4}$$

In order to show the unperturbed function to which the correction is being calculated, we introduce a second suffix in the coefficients a_k, writing

$$\Psi_n = \sum_n a_{kn}(t) \Psi_k^{(0)}.$$

Accordingly, we write the result of integrating equation (40.4) in the form

$$a_{kn}^{(1)} = -\frac{\mathrm{i}}{\hbar} \int V_{kn}(t) \mathrm{d}t = -\frac{\mathrm{i}}{\hbar} \int V_{kn} \mathrm{e}^{\mathrm{i}\omega_{kn}t} \mathrm{d}t. \tag{40.5}$$

This determines the wave functions in the first approximation.

Let us now consider in more detail the important case of a perturbation which is periodic with respect to time, of the form

$$\widehat{V} = \widehat{F} \mathrm{e}^{-\mathrm{i}\omega t} + \widehat{G} \mathrm{e}^{\mathrm{i}\omega t}, \tag{40.6}$$

where $\widehat{F}$ and $\widehat{G}$ are operators independent of time. Since V is Hermitian, we must have

$$\widehat{F} \mathrm{e}^{-\mathrm{i}\omega t} + \widehat{G} \mathrm{e}^{\mathrm{i}\omega t} = \widehat{F}^{\dagger} \mathrm{e}^{\mathrm{i}\omega t} + \widehat{G}^{\dagger} \mathrm{e}^{-\mathrm{i}\omega t}$$

whence $\widehat{G} = \widehat{F}^{\dagger}$, i.e.

$$G_{nm} = F_{mn}^* \tag{40.7}$$

This relation shows that

$$V_{kn}(t) = V_{kn} \mathrm{e}^{\mathrm{i}\omega_{kn}t} = F_{kn} \mathrm{e}^{\mathrm{i}(\omega_{kn} - \omega)t} + F_{nk}^* \mathrm{e}^{\mathrm{i}(\omega_{kn} + \omega)t}. \tag{40.8}$$

Substituting in (40.5) and integrating, we obtain the following expression for the expansion coefficients of the wave functions:

$$a_{kn}^{(1)} = -\frac{F_{kn}e^{i(\omega_{kn}-\omega)t}}{\hbar(\omega_{kn}-\omega)} - \frac{F_{kn}^* e^{i(\omega_{kn}+\omega)t}}{\hbar(\omega_{kn}+\omega)}. \tag{40.9}$$

These expressions are applicable if none of the denominators vanishes,[4] i.e. if for all k (and the given n)

$$E_k^{(0)} - E_n^{(0)} \neq \pm\hbar\omega \tag{40.10}$$

In a number of applications it is useful to have expressions for the matrix elements of an arbitrary quantity f, defined with respect to the perturbed wave functions. In the first approximation we have

$$f_{nm}(t) = f_{nm}^{(0)}(t) + f_{nm}^{(1)}(t),$$

where

$$f_{nm}^{(0)}(t) = \int \Psi_n^{(0)*}\widehat{f}\Psi_m^{(0)}\mathrm{d}q = f_{nm}^{(0)}e^{i\omega_{nm}t},$$

$$f_{nm}^{(1)}(t) = \int(\Psi_n^{(0)*}\widehat{f}\Psi_m^{(1)} + \Psi_n^{(1)*}\widehat{f}\Psi_m^{(0)})\mathrm{d}q.$$

Substituting here $\Psi_n^{(1)} = \sum_k a_{kn}^{(1)}\Psi_k^{(0)}$, with $a_{kn}^{(1)}$ determined by formula (40.9), it is easy to obtain the required expression

$$f_{nm}^{(1)}(t) = -e^{i\omega_{nm}t}\sum_k \left\{ \left[\frac{f_{nk}^{(0)}F_{km}}{\hbar(\omega_{km}-\omega)} + \frac{f_{km}^{(0)}F_{nk}}{\hbar(\omega_{kn}+\omega)}\right]e^{-i\omega t} + \right.$$
$$\left. + \left[\frac{f_{nk}^{(0)}F_{mk}^*}{\hbar(\omega_{km}+\omega)} + \frac{f_{km}^{(0)}F_{nk}^*}{\hbar(\omega_{kn}-\omega)}\right]e^{i\omega t}\right\}. \tag{40.11}$$

This formula is applicable if none of its terms becomes large, i.e. if none of the frequencies ω_{kn}, ω_{km} is too close to ω. For $\omega = 0$ we return to formula (38.12).

In all the formulae given here, it is understood that there is only a discrete spectrum of unperturbed energy levels. However, these formulae can be immediately generalized to the case where there is also a continuous spectrum (as before, we are concerned with the perturbation of states of the discrete spectrum); this is done by simply adding to the sums over the levels of the

---

[4] More precisely, if none is so small that the quantities $a_{kn}^{(1)}$ are no longer small compared with unity.

discrete spectrum the corresponding integrals over the continuous spectrum. Here it is necessary for the denominators $\omega_{kn} \pm \omega$ in formulae (40.9), (40.11) to be non-zero when the energy $E_k^{(0)}$ takes all values, not only of the discrete but also of the continuous spectrum. If, as usually happens, the continuous spectrum lies above all the levels of the discrete spectrum, then, for instance the condition (40.10) must be supplemented by the condition

$$E_{\min}^{(0)} - E_n^{(0)} > \hbar\omega \tag{40.12}$$

where $E_{\min}^{(0)}$ is the energy of the lowest level of the continuous spectrum.

PROBLEM

1. Determine the change in the nth and mth solutions of Schrödinger's equation in the presence of a periodic perturbation (of the form (40.6)), of frequency such that $E_m^{(0)} E_n^{(0)} = \hbar(\omega + \varepsilon)$, where ε is a small quantity.

SOLUTION. The method developed in the text is here inapplicable, since the coefficient $a_{mn}^{(1)}$ in (40.9) becomes large. We start afresh from the exact equations (40.3), with $V_{mk}^{(t)}$ given by (40.8). It is evident that the most important effect is due to those terms, in the sums on the right-hand side of equations (40.3), in which the time dependence is determined by the small frequency $\omega_{mn} - \omega$. Omitting all other terms, we obtain a system of two equations:

$$i\hbar\frac{\mathrm{d}a_m}{\mathrm{d}t} = F_{mn}e^{i(\omega_{mn}-\omega)t}a_n = F_{mn}e^{i\varepsilon t}a_n, \quad i\hbar\frac{a_n}{\mathrm{d}t} = F_{mn}^* a_m.$$

We make the substitution

$$a_n e^{i\varepsilon t} = b_n$$

and obtain the equations

$$i\hbar\dot{a}_m = F_{mn}b_n, \quad i\hbar(\dot{b}_n - i\varepsilon b_n) = F_{mn}^* a_m.$$

Eliminating a_m, we have

$$\ddot{b}_n - i\varepsilon\dot{b}_n + (1/\hbar^2)|F_{mn}|^2 b_n = 0.$$

We can take as two independent solutions of these equations

$$a_n = Ae^{i\alpha t}, \quad a_m = -A\frac{\hbar\alpha_1}{F_{mn}^*}e^{i\alpha_2 t} \tag{1}$$

and

$$a_n = Be^{-i\alpha_2 t}, \quad a_m = B\frac{\hbar\alpha_2}{F_{mn}^*}e^{-i\alpha_1 t}, \tag{2}$$

where A and B are constants (which have to be determined from the normalization condition), and we have used the notation

$$\alpha_1 = -\varepsilon/2 + \Omega, \quad \alpha_2 = \varepsilon/2 + \Omega, \quad \Omega = \sqrt{\varepsilon^2/4 + |\eta|^2}, \quad \eta = F_{mn}/\hbar.$$

Thus, under the action of the perturbation, the functions $\Psi_n^{(0)}, \Psi_m^{(0)}$ become $a_n \Psi_n^{(0)} + a_m \Psi_m^{(0)}$, with a_n and a_m given by (1) and (2).

Let the system be in the state $\Psi_m^{(0)}$ at the initial instant $(t = 0)$. The state of the system at subsequent instants is given by a linear combination of the two functions which we have obtained, which becomes $_m^{(0)}$ for $t = 0$:

$$\Psi = e^{i\varepsilon t/2} \left(\cos \Omega t - \frac{i\varepsilon}{2\Omega} \sin \Omega t \right) \Psi_m^{(0)} - \frac{i\eta^*}{\Omega} e^{-i\varepsilon t/2} \sin \Omega t \cdot \Psi_n^{(0)}. \qquad (40.13)$$

The squared modulus of the coefficient of $\Psi_n^{(0)}$ is

$$\frac{|\eta|^2}{2\Omega^2} \left[1 - \cos(2\Omega t) \right]. \qquad (4)$$

This gives the probability of finding the system in the state $\Psi_n^{(0)}$ at time t. We see that it is a periodic function with frequency 2Ω, and varies from 0 to $|\eta|^2/\Omega^2$.

For $\varepsilon = 0$(exact resonance) the probability (4) becomes

$$(1/2) \left[1 - \cos(2|\eta|t) \right].$$

It varies periodically between 0 and 1; in other words, the system makes periodic transitions from the state $\Psi_m^{(0)}$ to the state $\Psi_n^{(0)}$.

§ 41. Transitions under a perturbation acting for a finite time

Let us suppose that the perturbation $V(t)$ acts only during some finite interval of time (or that $V(t)$ diminishes sufficiently rapidly as $t \to \pm\infty$). Let the system be in the nth stationary state (of a discrete spectrum) before the perturbation begins to act (or in the limit as $t \to -\infty$). At any subsequent instant the state of the system will be determined by the function

$$\Psi = \sum_k a_{kn} \Psi_k^{(0)},$$

where, in the first approximation,

$$a_{kn} = a_{kn}^{(1)} = -\frac{i}{\hbar} \int_{-\infty}^{t} V_{kn} e^{i\omega_{kn}t} dt, \quad k \neq n,$$
$$a_{nn} = 1 + a_{nn}^{(1)} = 1 - \frac{i}{\hbar} \int_{-\infty}^{t} V_{nn} dt; \qquad (41.1)$$

the limits of integration in (40.5) are taken so that, as $t \to -\infty$, all the $a_{kn}^{(1)}$ tend to zero. After the perturbation has ceased to act (or in the limit

$t \to \infty$), the coefficients a_{kn} take constant values $a_{kn}(\infty)$, and the system is in the state with wave function

$$\Psi = \sum_k a_{kn}(\infty) \Psi_k^{(0)},$$

which again satisfies the unperturbed wave equation, but is different from the original function $\Psi_n^{(0)}$. According to the general rule, the squared modulus of the coefficient $a_{kn}(\infty)$ determines the probability for the system to have an energy $E_k^{(0)}$, i.e. to be in the kth stationary state.

Thus, under the action of the perturbation, the system may pass from its initial stationary state to any other. The probability of a transition from the initial (ith) to the final (fth) stationary state is[5]

$$w_{fi} = \frac{1}{\hbar^2} \left| \int_{-\infty}^{+\infty} V_{fi} e^{i\omega_{fi}t} dt \right|^2. \tag{41.2}$$

Let us now consider a perturbation which, once having begun, continues to act for an indefinite time (always, of course, remaining small). In other words, $V(t)$ tends to zero as $t - \infty$ and to a finite non-zero limit as $t \to +\infty$. Formula (41.2) cannot be applied directly here, since the integral in it diverges. This divergence, however, is physically unimportant and can easily be removed. To do this, we integrate by parts:

$$a_{fi} = -\frac{i}{\hbar} \int_{-\infty}^{t} V_{fi} e^{i\omega_{fi}t} dt - \frac{V_{fi} e^{i\omega_{fi}t}}{\hbar \omega_{fi}} \Big|_{-\infty}^{t} + \int_{-\infty}^{t} \frac{\partial V_{fi}}{\partial t} \frac{e^{i\omega_{fi}t}}{\hbar \omega_{fi}} dt.$$

The value of the first term vanishes at the lower limit, while at the upper limit it is formally identical with the expansion coefficients in formula (38.8); the presence of an additional periodic factor $e^{i\omega_{fi}t}$ is merely due to the fact that the a_{fi} are the expansion coefficients of the complete wave function Ψ, while the c_{fi} in §38 are the expansion coefficients of the time-independent function ψ. Hence it is clear that its limit as $t \to \infty$ gives simply the change in the original wave function $_i^{(0)}$ under the action of the "constant" part $V(+\infty)$ of the perturbation, and consequently has no relation to transitions into other states. The probability of a transition is given by the squared modulus of the second term and is

$$w_{fi} = \frac{1}{\hbar^2 \omega_{fi}^2} \left| \int_{-\infty}^{+\infty} \frac{\partial V_{fi}}{\partial t} e^{i\omega_{fi}t} dt \right|^2. \tag{41.3}$$

---

[5] For uniformity, the initial and final states will henceforward be denoted by i and f when transition probabilities are discussed. The suffixes of these probabilities will be written in the order fi, the same as for matrix elements.

The derivation is also valid when the transition is from a state of the discrete spectrum to a state of the continuous spectrum. The only difference is that here we have the probability of the transition from a given (ith) state to states in a range of values of ν_f (see the end of §38 from ν_f to $\nu_f + d\nu_f$, so that, for example, formula (41.2) must be written

$$dw_{if} = \frac{1}{\hbar^2} \left| \int_{-\infty}^{+\infty} V_{fi} e^{i\omega_{fi}t} dt \right|^2 d\nu_f. \tag{41.4}$$

If the perturbation $V(t)$ varies little during time intervals of the order of the period $\sim 1/\omega_{fi}$ the value of the integral in (41.2) or (41.3) will be very small. In the limit when the applied perturbation varies arbitrarily slowly, the probability of any transition with change of energy (i.e. with a non-zero frequency $1/\omega_{fi}$) tends to zero. Thus, when the applied perturbation changes sufficiently slowly (*adiabatically*), a system in any non-degenerate stationary state will remain in that state (see also §53).

In the opposite limiting case of a very rapid, "*instantaneous*" application of the perturbation, the derivatives $\partial V_{fi}/\partial t$ become infinite at the "instant of application". In the integral of $\frac{\partial V_{fi}}{\partial t} e^{i\omega_{fi}t}$, we can take outside the integral the comparatively slowly varying factor $e^{i\omega_{fi}t}$ and use its value at this instant. The integral is then found at once, and we obtain

$$w_{fi} = \frac{|V_{fi}|^2}{\hbar^2 \omega_{fi}^2} \tag{41.5}$$

The transition probabilities in instantaneous perturbations can also be found in cases where the perturbation is not small. Let the system be in a state described by one of the eigenfunctions $\psi_i^{(0)}$ of the original Hamiltonian $\widehat{H}_0$. If the change in the Hamiltonian occurs instantaneously (i.e. in a time short compared with the periods $1/\omega_{fi}$ of transitions from the given state i to other states), then the wave function of the system is "unable" to vary and remains the same as before the perturbation. It will no longer, however, be an eigenfunction of the new Hamiltonian $\widehat{H}$ of the system, i.e. the state $\psi_i^{(0)}$ will not be a stationary state. The probabilities w_{fi} for transitions of the system into the new stationary states are determined, according to the general rules of quantum mechanics, by the coefficients in the expansion of the function $\psi_i^{(0)}$ in terms of the eigenfunctions ψ_f of the Hamiltonian $\widehat{H}$:

$$w_{fi} = \left| \int \psi_i^{(0)} \psi_f^* dq \right|^2. \tag{41.6}$$

We shall show how this general formula becomes (41.5) if the change in the Hamiltonian $\widehat{V} = \widehat{H} - \widehat{H}_0$ is small. We multiply the equations

$$\widehat{H}_0 \psi_i^{(0)} = E_i^{(0)} \psi_i^{(0)}, \quad \widehat{H}^* \psi_f^* = E_f \psi_f^*$$

by ψ_f^* and $\psi_i^{(0)}$ respectively, integrate with respect to $\mathrm{d}q$ and subtract. Using also the self-conjugacy of the operator $\widehat{H}$, we obtain

$$(E_f - E_i^{(0)}) \int \psi_f^* \psi_i^{(0)} \mathrm{d}q = \int \psi_f^* \widehat{V} \psi_i^{(0)} \mathrm{d}q.$$

If the perturbation $\widehat{V}$ is small, in the first approximation we can replace E_f by the adjoining unperturbed level $E_f^{(0)}$, and the wave function ψ_f (on the right-hand side of the equation) by the corresponding function $\psi_f^{(0)}$. This gives

$$\int \psi_f^* \psi_i^{(0)} \mathrm{d}q = \frac{1}{\hbar \omega_{fi}} \int \psi_f^{(0)*} \widehat{V} \psi_i^{(0)} \mathrm{d}q,$$

and formula (41.6) becomes (41.5).

PROBLEMS

1. A uniform electric field is suddenly applied to a charged oscillator in the ground state. Determine the probabilities of transitions of the oscillator to excited states under the action of this perturbation.

SOLUTION. The potential energy of the oscillator in the uniform field (which exerts a force F on it) is

$$U(x) = \frac{m\omega^2}{2} x^2 - Fx \frac{m\omega^2}{2} (x - x_0)^2 + \text{const},$$

(where $x_0 = F/m\omega^2$), i.e. has still the pure oscillator form but with the equilibrium position shifted. Hence the wave functions of the stationary states of the perturbed oscillator are $\psi_k(x - x_0)$, where $\psi_k(x)$ are the oscillator functions (23.12); the initial wave function is $\psi_0(x)$ (23.13). Using these functions and the expression (23.11) for the Hermite polynomials, we find

$$\int_{-\infty}^{+\infty} \psi_0^{(0)} \psi_k \mathrm{d}x = \frac{(-1)^k}{\sqrt{2^k \pi k!}} e^{-\xi_0^2/2} \int_{-\infty}^{+\infty} e^{-\xi \xi_0} \frac{\mathrm{d}^k}{\mathrm{d}\xi^k} e^{-\xi^2 + 2\xi \xi_0} \mathrm{d}\xi,$$

with the notation $\xi_0 = x_0 \sqrt{m\omega/\hbar}$. On integrating k times by parts, the integral on the right becomes

$$\xi_0^k \int_{-\infty}^{+\infty} \exp(-\xi^2 + \xi \xi_0) \mathrm{d}\xi = \xi_0^k \sqrt{\pi} \exp \frac{\xi_0^2}{4}.$$

Thus the transition probability (41.6) is

$$w_{k0} = \frac{\overline{k}^k}{k!} \mathrm{e}^{-\overline{k}}, \quad \overline{k} = \frac{\xi_0^2}{2} = \frac{F^2}{2m\hbar\omega^3}.$$

As a function of the number k it represents a Poisson distribution for which the mean value of $\overline{k}$ is .

Perturbation theory is applicable when F is small, so that $\overline{k} \ll 1$. Then the excitation probabilities are small, and decrease rapidly with increasing $\overline{k}$. The largest is $w_{10} \approx \overline{k}$.

In the opposite case of large $F(\overline{k} \gg 1)$, excitation of the oscillator occurs with very high probability: the probability that the oscillator will remain in the normal state is $w_{00} = \mathrm{e}^{-\overline{k}}$.

2. The nucleus of an atom in the normal state receives an impulse which gives it a velocity v; the duration τ of the impulse is assumed short in comparison both with the electron periods and with a/v, where a is the dimension of the atom. Determine the probability of excitation of the atom under the influence of such a "jolt" (A. B. Migdal 1939).

SOLUTION. We use a frame of reference K' moving with the nucleus after the impact. By virtue of the condition $\tau \ll a/v$, the nucleus may be regarded as practically stationary during the impact, so that the coordinates of the electrons in K' and in the original frame K immediately after the perturbation are the same. The initial wave function in K' is

$$\psi_0' = \psi_0 \exp(-\mathrm{i}\boldsymbol{q} \sum_a \boldsymbol{r}_a), \quad \boldsymbol{q} = \frac{m\boldsymbol{v}}{\hbar},$$

where ψ_0 is the wave function of the normal state with the nucleus at rest, and the summation in the exponent is over all Z electrons in the atom. The required probability of transition to the kth excited state is now given, according to (41.6), by

$$w_{k0} = \left| \langle k | \exp(-\mathrm{i}\boldsymbol{q} \sum_a \boldsymbol{r}_a) | 0 \rangle \right|^2.$$

In particular, if $qa \ll 1$, then by expanding the exponential factor in the integrand and noting that the integral of $\psi_k^* \psi_0$ is zero because the functions ψ_0 and ψ_k are orthogonal, we obtain

$$w_{k0} = \left| \left\langle k \left| \boldsymbol{q} \sum_a \boldsymbol{r}_a \right| 0 \right\rangle \right|^2.$$

3. Determine the total probability of excitation and ionization of an atom of hydrogen which receives a sudden "jolt" (see Problem 2).

SOLUTION.

The required probability can be calculated as the difference

$$1 - w_{00} = 1 - \left| \int \psi_0^2 \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\boldsymbol{r}} \mathrm{d}V \right|,$$

where w_{00} is the probability that the atom will remain in the ground state ($\psi_0 = (\pi a^3)^{-1/2}e^{-r/a}$ being the wave function of the ground state of the hydrogen atom, with a the Bohr radius) Calculation of the integral gives

$$1 - w_{00} = 1 - \frac{1}{(1 + \frac{1}{4}q^2a^2)^4}.$$

In the limiting case $qa \ll 1$ this probability tends to zero as q^2a^2, while for $qa \gg 1$ it tends to unity as $1 - (2/qa)^8$.

4. Determine the probability that an electron will leave the K-shell of an atom with large atomic number Z when the nucleus undergoes β-decay. The velocity of the β-particle is assumed large in comparison with that of the K-electron (A. B. Migdal and E. L. Feinberg 1941).

SOLUTION.[6]

In the conditions stated the time taken by the β-particle to pass through the K-shell is small compared with the period of revolution of the electron, so that the change in the nuclear charge can be regarded as instantaneous. The perturbation is here represented by the change $V = 1/r$ in the field of the nucleus when the change in its charge is small (1 compared with Z). According to (41.5) the transition probability for one of the two K-shell electrons with energy $E_0 = -Z^2/2$ (here and below we use the fact that the state of the K-electrons is hydrogen-like; see §74) to a state of the continuous spectrum with energy $E = k^2/2$ in the range $dE = kdk$ is

$$dw = 2\frac{4|V_{0k}|^2}{(k^2 + Z^2)^2}dk$$

In the range which determines the matrix element V_{0k}, the important part is that of short distances ($\sim 1/Z$) from the nucleus, in which the hydrogen-like expression can again be used for the wave function of a state of the continuous spectrum. The final state of the electron must have angular momentum $l = 0$ (the same as that of the initial state). By means of the functions R_{l0}, and R_{k0} (normalized on the $k/2\pi$ scale), derived in §36 and formula (f.3) in the Mathematical Appendices we find[7]

$$\left(\frac{1}{r}\right)_{0k} = \frac{4\sqrt{2}\pi k}{1 - e^{-2\pi Z/k}}\frac{(1 + ik/Z)^{iZ/k}(1 - ik/Z)^{-iZ/k}}{1 + k^2/Z^2}$$

and, since

$$|(1 + i\alpha)^{i/\alpha}|^2 = \exp\left(-2\frac{\arctan\alpha}{\alpha}\right),$$

we obtain finally

$$dw = \frac{2^7}{Z^4(1 + k^2/Z^2)^4}f\left(\frac{k}{Z}\right)kdk,$$

---

[6] In Problems 4 and 5, atomic units are used.

[7] In the calculation it is convenient to use Coulomb units and then return to atomic units in the final result.

with

$$f(\alpha) = \frac{1}{1 - e^{-2\pi/\alpha}} \exp\left(-4\frac{\arctan\alpha}{\alpha}\right).$$

The limiting values of the function $f(\alpha)$ are e^{-4} for $\alpha \ll 1$ and $\alpha/2\pi$ for $\alpha \gg 1$.

The total probability of ionization of the K-shell is obtained by integration of dw over all energies of the emergent electron. A numerical evaluation gives $w = 0.65Z^{-2}$.

5. Determine the probability of emergence of an electron from the K-shell of an atom with large Z in α-decay of the nucleus. The velocity of the α-particle is small compared with that of the K-electron, but the time which it takes to leave the nucleus is small in comparison with the time of revolution of the electron (A. B. Migdal 1941, J. S. Levinger 1953).

SOLUTION. After the emergence of the α-particle, the perturbation acting on the electron is adiabatic The required effect is therefore determined essentially by the interval of time close to the "instant of application" of the perturbation which destroys the adiabaticity, when the α-particle, leaving the nucleus and moving freely, is still at a distance small compared with the radius of the K-orbit. The perturbation V which causes the ionization of the atom is here represented by the deviation of the combined field of the nucleus and the α-particle from the purely Coulomb field Z/r. The dipole moment of two particles with atomic weights 4 and $A - 4$, and charges 2 and $Z - 2$, at a distance vt apart (where v is the relative velocity of the nucleus and the α-particle), is

$$\frac{2(A - 4) - (Z - 2)4}{A}vt = \frac{2(A - 2Z)}{A}vt.$$

Hence the dipole term in the field of the nucleus and the α-particle is[8]

$$V = \frac{2(A - 2Z)}{A}vt\frac{z}{r^3},$$

where the z-axis is in the direction of the velocity $\mathbf{v}$. The matrix element of this perturbation reduces to that of z: taking the matrix element of the equation of motion of the electron $\ddot{z} = -Zz/r^3$, we obtain

$$\left(\frac{z}{r^3}\right)_{0k} = \frac{(E - E_0)^2}{Z}z_{0k}.$$

The required transition probability for one of the two electrons in the K-shell is, by (41.2),

$$dw = 2\left|\int_0^\infty V_{0k}e^{i(E_0 - E)t}dt\right|^2 dk = \frac{8(A - 2Z)^2v^2}{A^2Z^2}|z_{0k}|^2\frac{dk}{2\pi}$$

---

[8] If the difference $A - 2Z$ is small, it may be necessary to take account of the next (quadrupole) term also.

to calculate the integral, we include in the integrand an additional damping factor $e^{-\lambda t}$ with $\lambda > 0$, and then make $\lambda \to 0$ in the result. To calculate the matrix element of $z = r \cos \theta$, we note that, since the orbital angular momentum in the initial state is $l = 0$, $\cos \theta$ has a non-zero matrix element only for the transition to a state with $l = 1$, and

$$|(\cos \theta)_{01}|^2 = \frac{1}{3},$$

and

$$|z_{0k}|^2 = \frac{1}{3}|r_{0k}|^2.$$

Calculating r_{0k} by means of the radial functions R_{00} and R_{k1}, we find

$$\mathrm{d}w = \frac{2^{11}(A - 2Z)^2 v^2}{3 A^2 Z^6 (1 + k^2/Z^2)^5} f\left(\frac{k}{Z}\right) k \mathrm{d}k$$

the function f being as in Problem 4.

§ 42. Transitions under the action of a periodic perturbation

The results are different for the probability of transitions to the states of the continuous spectrum under the action of a periodic perturbation. Let us suppose that, at some initial instant $t = 0$, the system is in the ith stationary state of the discrete spectrum. We shall assume that the frequency ω of the periodic perturbation is such that

$$\hbar\omega > E_{\min} - E - i^{(0)} \tag{42.1}$$

where $E_{\min}$ is the value of the energy where the continuous spectrum begins.

It is evident from the results of §40 that the chief part will be played by states of the continuous spectrum with energies E_f very close to the resonance energy $E_i^{(0)} + \hbar\omega$, i.e. those for which the difference $\omega_{fi} - \omega$ is small. For this reason it is sufficient to consider, in the matrix elements (40.8) of the perturbation, only the first term (with the frequency $\omega_{fi} - \omega$ close to zero). Substituting this term in (40.5) and integrating, we obtain

$$a_{fi} = -\frac{i}{\hbar} \int_0^t V_{fi}(t) \mathrm{d}t = -F_{fi} \frac{\exp\left[i(\omega_{fi} - \omega)t\right] - 1}{\hbar(\omega_{fi} - \omega)} \tag{42.2}$$

The lower limit of integration is chosen so that $a_{fi} = 0$ for $t = 0$, in accordance with the initial condition imposed.

Hence we find for the squared modulus of a_{fi}

$$|a_{fi}|^2 = |F_{fi}|^2 \frac{4 \sin^2 \frac{\omega_{fi}-\omega}{2}t}{\hbar^2(\omega_{fi}-\omega)^2}. \tag{42.3}$$

It is easy to see that, for large t, this function can be regarded as proportional to t. To show this, we notice that

$$\lim_{t\to\infty} \frac{\sin^2 \alpha t}{\pi t \alpha^2} = \delta(\alpha) \tag{42.4}$$

For when a $\alpha \neq 0$ this limit is zero, while for $\alpha = 0$ we have $\frac{\sin^2 \alpha t}{t\alpha^2} = t$, so that the limit is infinite; finally, integrating over α from $-\infty$ to $+\infty$, we have (with the substitution $\alpha t = \xi$)

$$\frac{1}{\pi} = \int_{-\infty}^{+\infty} \frac{\sin^2 \alpha t}{t\alpha^2}d\alpha = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\sin^2 \xi}{\xi^2}d\xi = 1.$$

Thus the function on the left-hand side of equation (42.4) in fact satisfies all the conditions which define the delta function. Accordingly, we can write for large t

$$|a_{fi}|^2 = \frac{1}{\hbar^2}|F_{fi}|^2 \pi t \delta\left(\frac{\omega_{fi}-\omega}{2}\right)$$

or, substituting $\hbar\omega_{fi} = E_f - E_i^{(0)}$ and using the fact that $\delta(ax) = \delta(x)/\alpha$:

$$|a_{fi}|^2 = \frac{2\pi}{\hbar}|F_{fi}|^2 \delta\left(E_f - E_i^{(0)} - \hbar\omega\right)t.$$

The expression $|a_{fi}|^2 d\nu_f$ is the probability of a transition from the original state to one in the interval $d\nu_f$. We see that, for large t, it is proportional to the time interval elapsed since $t = 0$. The probability dw_{fi} of the transition per unit time is[9]

$$dw_{fi} = \frac{2\pi}{\hbar}|F_{fi}|^2 \delta\left(E_f - E_i^{(0)} - \hbar\omega\right)d\nu_f. \tag{42.5}$$

As we should expect, it is zero except for transitions to states with energy $E_f = E_i^{(0)} + \hbar\omega$. If the energy levels of the continuous spectrum are not degenerate, so that ν_f can be taken as the value of the energy alone, then

---

[9] It is easy to verify that, on taking account of the second term in (40.8), which we have omitted, additional expressions are obtained which, on being divided by t, tend to zero as $t \to +\infty$.

the whole "interval" of states $\mathrm{d}\nu_f$ reduces to a single state with energy $E = E_i^{(0)} + \hbar\omega$, and the probability of a transition to this state is

$$w_{Ei} = \frac{2\pi}{\hbar}|F_{Ei}|^2. \tag{42.6}$$

There is another method of deriving formula (42.5) that is methodologically instructive, in which the periodic perturbation is assumed not to be applied at a time $t = 0$ but to increase slowly from $t = -\infty$ by an exponential law $\mathrm{e}^{\lambda t}$ with a positive constant λ which is then made to tend to zero (*adiabatic switch-on*). The initial condition $a_{fi} = 0$ is accordingly applied at $t = -\infty$. The matrix element of the perturbation now has the form

$$V_{fi} = F_{fi}\mathrm{e}^{\mathrm{i}(\omega_{fi}-\omega)t+\lambda t},$$

and (42.2) becomes

$$a_{fi} = -\frac{\mathrm{i}}{\hbar}\int_{-\infty}^{t} V_{fi}(t)\mathrm{d}t - F_{fi}\frac{\exp\left[\mathrm{i}(\omega_{fi}-\omega)t+\lambda t\right]}{\hbar(\omega_{fi}-\omega-\mathrm{i}\lambda)}. \tag{42.7}$$

Hence

$$|a_{fi}|^2 = \frac{1}{\hbar^2}|F_{fi}|^2\frac{\mathrm{e}^{2\lambda t}}{(\omega_{fi}-\omega)^2+\lambda^2}.$$

The transition probability per unit time is given by the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}|a_{fi}|^2 = 2\lambda|a_{fi}|^2.$$

There is a formula

$$\lim_{\lambda\to 0}\frac{\lambda}{\pi(\alpha^2+\lambda^2)} = \delta(\alpha), \tag{42.8}$$

valid in the same sense as (42.4); with this we find, taking the limit $\lambda \to 0$:

$$\frac{\mathrm{d}}{\mathrm{d}t}|a_{fi}|^2 \to \frac{2\pi}{\hbar^2}|F_{fi}|^2\delta(\omega_{fi}-\omega),$$

and thus return to (42.5).

§ 43. Transitions in the continuous spectrum

One of the most important applications of perturbation theory is to calculate the probability of a transition in the continuous spectrum under the

action of a constant (time-independent) perturbation. We have already mentioned that the states of the continuous spectrum are almost always degenerate. Having chosen in some manner the set of unperturbed wave functions corresponding to some given energy level, we can put the problem as follows. It is known that, at the initial instant, the system is in one of these states; it is required to determine the probability of the transition to another state with the same energy. For transitions from the initial state i to states between ν_f and $\nu_f + \mathrm{d}\nu_f$ we have at once from (42.5) (putting $\omega = 0$ and changing the notation)

$$\mathrm{d}w_{fi} = \frac{2\pi}{\hbar}|V_{fi}|^2\delta\left(E_f - E_i\right)\mathrm{d}\nu_f. \tag{43.1}$$

This expression is, as we should expect, zero except for $E_f = E_i$: under the action of a constant perturbation, transitions occur only between states with the same energy. It must be noticed that, for transitions from states of the continuous spectrum, the quantity $\mathrm{d}w_{fi}$ cannot be regarded directly as the transition probability; it is not even of the right dimensions (1/time). Formula (43.1) represents the number of transitions per unit time, and its dimensions depend on the chosen method of normalization of the wave functions of the continuous spectrum.[10]

Let us calculate the perturbed wave function, which before the action of the perturbation is the same as the original unperturbed function $\psi_i^{(0)}$. Using the method given at the end of §42, we can regard the perturbation as being adiabatically applied according to $\mathrm{e}^{\lambda t}$ with $\lambda \to 0$. From (42.7), putting $\omega = 0$ and changing the notation, we have

$$a_{fi}^1 = V_{fi}\frac{\exp\left\{\frac{\mathrm{i}}{\hbar}(E_f - E_i)t + \lambda t\right\}}{E_i - E_f + \mathrm{i}\lambda}. \tag{43.2}$$

The perturbed wave function is

$$\Psi_i = \Psi_i^{(0)} + \int a_{fi}^{(1)}\Psi_f^{(0)}\mathrm{d}\nu_f,$$
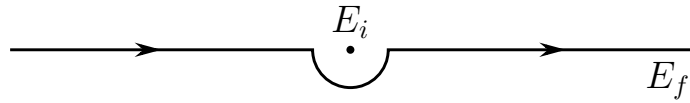
where the integration is extended over the whole of the continuous spectrum.[11] Substitution of (43.2) gives

$$\Psi_i = \left[\psi_i^{(0)} + \int V_{fi}\psi_f^{(0)}\frac{\mathrm{d}\nu_f}{E_i - E_f + \mathrm{i}0}\right]\exp\left(-\frac{\mathrm{i}}{\hbar}E_i t\right). \tag{43.3}$$

---

[10] The phenomena comprised within the theory here discussed include, for example, various types of collision; the system in its initial and final states is a set of free particles and the perturbation is the interaction between them. With appropriate normalization of the wave functions, (43.1) may then be the collision cross-section (see §126).

[11] If there is also a discrete spectrum, then we must add to the integral in this formula (and subsequent ones) the appropriate sum over the states of the discrete spectrum.

In the limit as $\lambda \to 0$, the factor $e^{\lambda t}$ becomes unity. The term $+i0$, denoting the limit of $i\lambda$ as λ tends to zero from positive values, determines the manner of integration with respect to the variable E_f (dE_f occurs as a factor in $d\nu_f$ together with the differentials of other quantities which describe the states of the continuous spectrum). Without the term $i\lambda$, the integrand in (43.3) would have a pole at $E_f = E_i$, near which the integral would diverge. The term $i\lambda$ moves this pole into the upper half-plane of the complex variable E_f. After the limit $\lambda \to 0$ is taken, the pole returns to the real axis, but we know that the path of integration must pass beneath it:

$$ \tag{43.4} $$

The time factor in (43.3) shows that this function belongs, as it should, to the same energy E_i as the original unperturbed function. In other words, the function

$$\psi_i = \psi_i^{(0)} + \int \frac{V_{fi}}{E_i - E_f + i0}\psi_f^{(0)} d\nu_f \tag{43.5}$$

satisfies Schrödinger's equation

$$(\widehat{H}_0 + \widehat{V})\psi_i = E_i\psi_i$$

It is therefore natural that the expression obtained should correspond exactly to (38.8).[12]

The calculations given above correspond to the first approximation of perturbation theory. It is not difficult to calculate the second approximation as well. To do this, we must derive the formula for the next approximation to Ψ_i; this is easily effected by using the method of §38 (now that we know the method of dealing with the "divergent" integrals). A simple calculation gives the formula

$$\Psi_i = \left\{\psi_i^{(0)} + \int\left[V_{fi} + \int\frac{V_{f\nu}V_{\nu i}}{E_i - E_f + i0}d\nu\right]\frac{\psi_f^{(0)}d\nu_f}{E_i - E_f + i0}\right\}\exp\left(-\frac{i}{\hbar}E_i t\right).$$

Comparing this expression with formula (43.3), we can write down the corresponding formula for the probability (or, more precisely, the number) of transitions, by direct analogy with (43.1):

$$dw_{fi} = \frac{2\pi}{\hbar}\left|V_{fi} + \int\frac{V_{f\nu}V_{\nu i}}{E_i - E_\nu + i0}\right|^2 \delta(E_i - E_f)d\nu_f. \tag{43.6}$$

---

[12] With this formula, the way in which the integral is to be taken can be found from the condition that the asymptotic expression for ψ_i at large distances should contain only an outgoing (and not an ingoing) wave (see §136).

It may happen that the matrix element V_{fi} for the transition considered vanishes. The effect is then zero in the first approximation, and (43.6) becomes

$$\mathrm{d}w_{fi} = \frac{2\pi}{\hbar} \left| \int \frac{V_{f\nu}V_{\nu i}}{E_i - E_\nu} \mathrm{d}\nu \right|^2 \delta(E_f - E_i)\mathrm{d}\nu_f, \qquad (43.7)$$

In applications of this formula, the point $E_\nu = E_i$ is not usually a pole of the integrand; the manner of integrating with respect to E_ν is then unimportant, and the integral can be taken along the real axis.

The states ν for which $V_{f\nu}$ and $V_{\nu i}$ are not zero are usually called *intermediate states* for the transition $i \to f$. Intuitively, we may say that this transition takes place as if in two steps $i \to \nu$ and $\nu \to f$ (but such a description must not be taken literally, of course). It may happen that the transition $i \to f$ can take place not through one but only through several successive intermediate states. Formula (43.7) can be at once generalized to such cases. For example, if two intermediate states are needed, we have

$$\mathrm{d}w_{fi} = \frac{2\pi}{\hbar} \left| \int \frac{V_{f\nu'}V_{\nu'\nu}V_{\nu i}}{(E_i - E_{\nu'})(E_i - E_\nu)} \mathrm{d}\nu\mathrm{d}\nu' \right|^2 \delta(E_f - E_i)\mathrm{d}\nu_f. \qquad (43.8)$$

Lastly, to clarify the mathematical significance of the integrals taken along a path of the form (43.4), we shall prove the formula

$$\int \frac{f(x)\mathrm{d}x}{x - a - \mathrm{i}0} = P \int \frac{f(x)\mathrm{d}x}{x - a} + \mathrm{i}\pi f(a), \qquad (43.9)$$

where the integration is along a segment of the real axis including the point $x = a$. If we pass round the pole $x = a$ along a semicircle of radius ρ, we find that the whole integral is equal to the sum of the integrals along the real axis from the lower limit to $a - \rho$ and from $a + \rho$ to the upper limit, together with $\mathrm{i}\pi$ times the residue of the integrand at the pole. In the limit $\rho \to 0$, the integrals along the real axis make the integral along the complete segment, taken as a principal value (denoted by P), and the result is (43.9), which may also be symbolically written

$$\frac{1}{x - a - \mathrm{i}0} = P\frac{1}{x - a} + \mathrm{i}\pi\delta(x - a); \qquad (43.10)$$

P here denotes the taking of the principal value when integrating the function $f(x)/(x - a)$.

§ 44. The uncertainty relation for energy

Let us consider a system composed of two weakly interacting parts. We suppose that it is known that at some instant these parts have definite values

of the energy, which we denote by E and ε respectively. Let the energy be measured again after some time interval Δt; the values E', ε' obtained are in general different from E, ε. It is easy to determine the order of magnitude of the most probable value of the difference $E' + \varepsilon' - E - \varepsilon$ which is found as a result of the measurement.

According to formula (42.3) with $\omega = 0$, the probability of a transition of the system (after time t), under the action of a time-independent perturbation, from a state with energy E to one with energy E' is proportional to

$$\left(\sin^2 \frac{E' - E}{2\hbar} t\right) / (E' - E)2.$$

Hence we see that the most probable value of the difference $E' - E$ is of the order of $\hbar/t$.

Applying this result to the case we are considering (the perturbation being the interaction between the parts of the system), we obtain the relation

$$|E_\varepsilon - E' - \varepsilon'|\Delta t \sim \hbar. \tag{44.1}$$

Thus the smaller the time interval Δt, the greater the energy change that is observed. It is important to notice that its order of magnitude $\hbar/\Delta t$ is independent of the amount of the perturbation. The energy change determined by the relation (44.1) will be observed, however weak the interaction between the two parts of the system. This result is peculiar to quantum theory and has a deep physical significance. It shows that, in quantum mechanics, the law of conservation of energy can be verified by means of two measurements only to an accuracy of the order of $\hbar/\Delta t$, where Δt is the time interval between the measurements.

The relation (44.1) is often called the uncertainty relation for energy. However, it must be emphasized that its significance is entirely different from that of the uncertainty relation $\Delta p \Delta x \sim \hbar$ for the coordinate and momentum. In the latter, Δp and Δx are the uncertainties in the values of the momentum and coordinate at the same instant; they show that these two quantities can never have entirely definite values simultaneously. The energies E, ε, on the other hand, can be measured to any degree of accuracy at any instant. The quantity $(E + \varepsilon) - (E' + \varepsilon')$ in (44.1) is the difference between two exactly measured values of the energy $E + \varepsilon$ at two different instants, and not the uncertainty in the value of the energy at a given instant.

If we regard E as the energy of some system and ε as that of a "measuring apparatus", we can say that the energy of interaction between them can be taken into account only to within $\hbar/\Delta t$. Let us denote by $\Delta E, \Delta \varepsilon, \ldots$ the errors in the measurements of the corresponding quantities. In the favourable

case when $\varepsilon, \varepsilon'$ are known exactly ($\Delta\varepsilon = \Delta\varepsilon' = 0$), we have

$$\Delta(E - E') \sim \frac{\hbar}{\Delta t}. \tag{44.2}$$

From this relation we can derive important consequences concerning the measurement of momentum. The process of measuring the momentum of a particle (for definiteness, we shall speak of an electron) consists in a collision of the electron with some other ("measuring") particle, whose momenta before and after the collision can be regarded as known exactly.[13]   If we apply to this collision the law of conservation of momentum, we obtain three equations (the three components of a single vector equation) in six unknowns (the components of the momentum of the electron before and after the collision). The number of equations can be increased by bringing about a series of further collisions between the electron and "measuring" particles, and applying to each collision the law of conservation of momentum. This, however, increases the number of unknowns also (the momenta of the electron between collisions), and it is easy to see that, whatever the number of collisions, the number of unknowns will always be three more than the number of equations. Hence, in order to measure the momentum of the electron, it is necessary to bring in the law of conservation of energy at each collision, as well as that of momentum. The former, however, can be applied, as we have seen, only to an accuracy of the order of $\hbar/\Delta t$, where Δt is the time between the beginning and end of the process in question.

To simplify the subsequent discussion, it is convenient to consider an imaginary idealized experiment in which the "measuring particle" is a perfectly reflecting plane mirror; only one momentum component is then of importance, namely that perpendicular to the plane of the mirror. To determine the momentum P of the particle, the laws of conservation of momentum and energy give the equations

$$p' + P' - p - P' = 0, \tag{44.3}$$

$$|\varepsilon' + E' - \varepsilon - E| \sim \frac{\hbar}{\Delta t} \tag{44.4}$$

where P, E are the momentum and energy of the particle, and p, ε those of the mirror; the unprimed and primed quantities refer to the instants before and after the collision respectively. The quantities $p, p', \varepsilon, \varepsilon'$ relating to the "measuring particle" can be regarded as known exactly, i.e. the errors in

---

[13] In the present analysis it is of no importance how the energy of the "measuring" particle is ascertained.

them are zero. Then we have for the errors in the remaining quantities, from the above equations:

$$\Delta P = \Delta P', \quad |\Delta E' - \Delta E| \sim \frac{\hbar}{\Delta t}.$$

But

$$\Delta E = (\partial E/\partial P)\Delta P = v\Delta P,$$

where v is the velocity of the electron (before the collision), and similarly

$$\Delta E' = v'\Delta P' = v'\Delta P.$$

Hence we obtain

$$|(v'_x - v_x)\delta P_x| \sim \frac{\hbar}{\delta t}. \tag{44.5}$$

We have here added the suffix x to the velocity and momentum, in order to emphasize that this relation holds for each of their components separately.

This is the required relation. It shows that the measurement of the momentum of the electron (with a given degree of accuracy ΔP) necessarily involves a change in its velocity (i.e. in the momentum itself). This change becomes greater as the duration of the measuring process becomes shorter. The change in velocity can be made arbitrarily small only as $\Delta t \to \infty$, but measurements of momentum occupying a long time can be significant only for a free particle. The non-repeatability of a measurement of momentum after short intervals of time, and the "two-faced" nature of measurement in quantum mechanics—the necessity of a distinction between the measured value of a quantity and the value resulting from the process of measurement—are here exhibited with particular clarity.[14]

The conclusion reached at the beginning of this section, which was based on perturbation theory, can also be derived from another standpoint by considering the decay of a system under the action of some perturbation. Let E_0 be some energy level of the system, calculated without any allowance for the possibility of its decay. We denote by τ the *lifetime* of this state of the system, i.e. the reciprocal of the probability of decay per unit time. Then we find by the same method that

$$|E_0 - E - \varepsilon| \sim \hbar/\tau \tag{44.6}$$

where E, ε are the energies of the two parts into which the system decays. The sum $E + \varepsilon$, however, gives us an estimate of the energy of the system

---

[14] The relation (44.5) and the elucidation of the physical significance of the uncertainty relation for energy are due to N. Bohr (1928).

before it decays. Hence the above relation shows that the energy of a system, in some "quasi-stationary" state, which is free to decay can be determined only to within a quantity of the order of $\hbar/\tau$. This quantity is usually called the *width* Γ of the level. Thus

$$\Gamma \sim \hbar/\tau \tag{44.7}$$

§ 45. Potential energy as a perturbation

The case where the total potential energy of the particle in an external field can be regarded as a perturbation merits special consideration. The unperturbed Schrödinger's equation is then the equation of free motion of the particle:

$$\Delta\psi^{(0)} + k^2\psi^{(0)} = 0, \quad k = \frac{\sqrt{2mE}}{\hbar} = \frac{p}{\hbar} \tag{45.1}$$

and has solutions which represent plane waves. The energy spectrum of free motion is continuous, so that we are concerned with an unusual case of perturbation theory in a continuous spectrum. The solution of the problem is here more conveniently obtained directly, without having recourse to general formulae.

The equation for the correction $\psi^{(1)}$ to the wave function in the first approximation is

$$\Delta\psi^{(1)} + k^2\psi^{(1)} = \frac{2mU}{\hbar^2}\psi^{(0)} \tag{45.2}$$

where U is the potential energy. The solution of this equation, as we know from electrodynamics, can be written in the form of *retarded potentials*, i.e. in the form[15]

$$\psi^{(1)} = -\frac{m}{2\pi\hbar^2}\int \psi^{(0)}U(x', y', z')\mathrm{e}^{\mathrm{i}kr}\frac{\mathrm{d}V'}{r}, \tag{45.3}$$

where

$$\mathrm{d}V' = \mathrm{d}x'\mathrm{d}y'\mathrm{d}z', \quad r^2 = (x - x')^2 + (y - y')^2 + (z - z')^2.$$

Let us find what conditions must be satisfied by the field U in order that it may be regarded as a perturbation. The condition of applicability of perturbation theory is contained in the requirement that $\psi^{(1)} \ll \psi^{(0)}$. Let a be the order of magnitude of the dimensions of the region of space

---

[15] This is a particular integral of equation (45.2), to which we may add any solution of the same equation with zero on the right-hand side, i.e. the unperturbed equation (45.1).

in which the field is noticeably different from zero. We shall first suppose that the energy of the particle is so small that ka is at most of the order of unity. Then the factor e^{ikr} in the integrand of (45.3) is unimportant in an order-of-magnitude estimate, and the integral is of the order of $\psi^{(0)}|U|a^2$, so that

$$\psi^{(1)} \sim (ma^2|U|\hbar^2)\psi^{(0)}$$

and we have the condition

$$|U| \ll \frac{\hbar^2}{ma^2} \qquad (\text{for} \quad ka \lesssim 1). \qquad (45.4)$$

We notice that the expression on the right has a simple physical meaning; it is the order of magnitude of the kinetic energy which the particle would have if enclosed in a volume of linear dimensions a (since, by the uncertainty relation, its momentum would be of the order of $\hbar/a$).

Let us consider, in particular, a potential well so shallow that the condition (45.4) holds for it. It is easy to see that in such a well there are no negative energy levels (R. Peierls 1929); this has been shown, for the particular case of a spherically symmetric well, in §33, Problem, For, when $E = 0$, the unperturbed wave function reduces to a constant, which can be arbitrarily taken as unity: $\psi^{(0)} = 1$. Since $\psi^{(1)} \ll \psi^{(0)}$, it is clear that the wave function $\psi = 1 + \psi^{(1)}$ for motion in the well nowhere vanishes; the eigenfunction, being without nodes, belongs to the normal state, so that $E = 0$ remains the least possible value of the energy of the particle. Thus, if the well is sufficiently shallow, only an infinite motion of the particle is possible: the particle cannot be "captured" by the well. Note that this result is peculiar to quantum theory; in classical mechanics a particle can execute a finite motion in any potential well.

It must be emphasized that all that has been said refers only to a three-dimensional well. In a one- or two-dimensional well (i.e. one in which the field is a function of only one or two coordinates), there are always negative energy levels (see the Problems at the end of this section). This is related to the fact that, in the one- and two-dimensional cases, the perturbation theory under consideration is inapplicable for an energy E which is zero (or very small).[16]

---

[16] In the two-dimensional case $\psi^{(1)}$ is expressed (as is known from the theory of the two-dimensional wave equation) as an integral similar to (45.3), in which, instead of $\frac{e^{ikr}}{r}dx'dy'dz'$ we have $i\pi H_0^{(1)}(kr)dx'dy'$, where $H_0^{(1)}$ is the Hankel function and $r = \sqrt{(x-x)^2 + (y-y)^2}$. As $k \to 0$, the Hankel function, and therefore the whole integral, tend logarithmically to infinity.

Similarly, in the one-dimensional case, we have, in the integrand, $2\pi i \frac{e^{ikr}}{k}dx'$, where $r = |x - x'|$, and as $k \to 0$ $\psi^{(1)}$ tends to infinity as $1/k$.

For large energies, when $ka \gg 1$, the factor e^{ikr} in the integrand plays an important part, and markedly reduces the value of the integral. The solution (45.3) in this case can be transformed; the alternative form, however, is more conveniently derived by returning to equation (45.2). We take as x-axis the direction of the unperturbed motion; the unperturbed wave function then has the form $\psi^{(0)} = \mathrm{e}^{ikx}$ (the constant factor is arbitrarily taken as unity). Let us seek a solution of the equation

$$\Delta\psi^{(1)} + k^2\psi^{(1)} = \frac{2m}{\hbar^2}U\mathrm{e}^{ikx}$$

in the form $\psi^{(1)} = \mathrm{e}^{ikx}f$; in view of the assumed large value of k, it is sufficient to retain in $\Delta\psi^{(1)}$ only those terms in which the factor e^{ikx} is differentiated one or more times. We then obtain for f the equation

$$2ik\frac{\partial f}{\partial x} = \frac{2mU}{\hbar^2},$$

whence

$$\psi^{(1)} = \mathrm{e}^{ikx}f = -\frac{im}{\hbar^2 k}\mathrm{e}^{ikx}\int U\,\mathrm{d}x. \tag{45.5}$$

An estimation of this integral gives $|\psi^{(1)}| \sim m|U|a/\hbar^2 k$, so that the condition of applicability of perturbation theory in this case is

$$|U| \ll \frac{\hbar^2}{ma^2}ka = \frac{\hbar v}{a}, \quad ka \gg 1. \tag{45.6}$$

where $v = k\hbar/m$ is the velocity of the particle. It is to be observed that this condition is weaker than (45.4). Hence, if the field can be regarded as a perturbation at small energies of the particle, it can always be so regarded at large energies, whereas the converse is not necessarily true.[17]

The applicability of the perturbation theory developed here to a Coulomb field requires special consideration. In a field where $U = \alpha/r$, it is impossible to separate a finite region of space outside which U is considerably less than inside it. The required condition can be obtained by writing in (45.6) a variable distance r instead of the parameter a; this leads to the inequality

$$\frac{\alpha}{\hbar v} \ll 1. \tag{45.7}$$

---

[17] In the one-dimensional case the condition for perturbation theory to be applicable is given by the inequality (45.6) for all ka. The derivation of the condition (45.4) given above for the three-dimensional case is not valid in the one-dimensional case, owing to the divergence of the resulting function $\psi^{(1)}$ (see the preceding footnote).

Thus, for large energies of the particle, a Coulomb field can be regarded as a perturbation.[18]

Finally, we shall derive a formula which approximately determines the wave function of a particle whose energy E everywhere considerably exceeds the potential energy U (no other conditions being imposed). In the first approximation, the wave function depends on the coordinates in the same way as for free motion (whose direction is taken as the x-axis). Accordingly, let us look for ψ in the form $\psi = \mathrm{e}^{\mathrm{i}kx}F$, where F is a function of the coordinates which varies slowly in comparison with the factor $\mathrm{e}^{\mathrm{i}kx}$ (but we cannot in general say that it is close to unity). Substituting in Schrödinger's equation, we obtain for F the equation

$$2\mathrm{i}k\frac{\partial F}{\partial x} = \frac{2m}{\hbar^2}UF, \tag{45.8}$$

whence

$$\psi = \mathrm{e}^{\mathrm{i}kx}F = \mathrm{const} \cdot \mathrm{e}^{\mathrm{i}kx}\exp\left(-\frac{\mathrm{i}}{\hbar v}\int U\,\mathrm{d}x\right). \tag{45.9}$$

This is the required expression. It should, however, be borne in mind that this formula is not valid at large distances. In equation (45.8) a term ΔF has been omitted which contains second derivatives of F. The derivative $\partial^2 F/\partial x^2$, together with the first derivative $\partial F/\partial x$, tends to zero at large distances, but the derivatives with respect to the transverse coordinates y and z do not tend to zero, and can be neglected only if $x \ll ka^2$.

PROBLEMS

1. Determine the energy level in a one-dimensional potential well whose depth is small. It is assumed that the condition (45.4) is satisfied.

SOLUTION. We make the hypothesis, which will be confirmed by the result, that the energy level $|E| \ll |U|$. Then, on the right-hand side of Schrödinger's equation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \frac{2m}{\hbar^2}(U(x) - E)\psi$$

we can neglect E in the region of the well, and regard ψ as a constant, which without loss of generality can be taken as unity:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \frac{2m}{\hbar^2}U.$$

We integrate this equation with respect to x between two points $\pm x_1$ such that $a \ll x_1 \ll 1/\varkappa$, where a is the width of the well and $\varkappa = \sqrt{2m|E|}/\hbar$. Since the

---

[18] It must be borne in mind that the integral (45.5) with a field $U = \alpha/r$ diverges (logarithmically) when $x/\sqrt{y^2 + z^2}$ is large. Hence the wave function in a Coulomb field, obtained by means of perturbation theory, is inapplicable within a narrow cone about the x-axis.

integral of $U(x)$ converges, the integration on the right can be extended to the whole range from $-\infty$ to $+\infty$:

$$\frac{\mathrm{d}\psi}{\mathrm{d}x}\bigg|_{-x_1}^{x_1} = \frac{2m}{\hbar^2} \int_{-\infty}^{+\infty} U\mathrm{d}x. \tag{1}$$

At large distances from the well, the wave function is of the form $\psi = \mathrm{e}^{\pm\varkappa x}$. Substituting this in (1), we find

$$-2\varkappa = \frac{2m}{\hbar^2} \int_{-\infty}^{+\infty} U\mathrm{d}x$$

or

$$|E| = \frac{m}{2\hbar^2} \left(\int_{-\infty}^{+\infty} U\mathrm{d}x \right)^2.$$

We see that, in accordance with the hypothesis, the energy of the level is a small quantity of a higher order (the second) than the depth of the well.

2. Determine the energy level in a two-dimensional potential well $U(r)$ (where r is the polar coordinate in the plane) of small depth; it is assumed that the integral $\int_0^\infty rU\mathrm{d}r$ converges.

SOLUTION. Proceeding as in the previous problem, we have in the region of the well the equation

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}\psi}{\mathrm{d}r} \right) = \frac{2m}{\hbar^2}U.$$

Integrating this with respect to r from 0 to r_1 (where $a \ll r_1 \ll 1/\varkappa$), we find

$$\frac{\mathrm{d}\psi}{\mathrm{d}r}\bigg|_{r=r_1} = \frac{2m}{\hbar^2 r_1} \int_0^\infty rU(r)\mathrm{d}r. \tag{1}$$

At large distances from the well, the equation of free motion in two dimensions is

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}\psi}{\mathrm{d}r} \right) + \frac{2m}{\hbar^2}E\psi = 0$$

and has a solution (vanishing at infinity) $\psi = \mathrm{const} \cdot H_0^{(1)}(\mathrm{i}\varkappa r)$; for small values of the argument, the leading term in this function is proportional to $\log \varkappa r$. Bearing this in mind, we equate the logarithmic derivatives of ψ for $r \sim a$ inside the well (the right-hand side of (1) and outside it, obtaining

$$\frac{1}{a\log\varkappa a} \approx \frac{2m}{\hbar^2 a} \int_0^\infty U(r)r\mathrm{d}r,$$

whence

$$|E| \sim \frac{\hbar^2}{ma^2} \exp\left[-\frac{\hbar^2}{m}\left| \int_0^\infty Ur\mathrm{d}r \right|^{-1} \right].$$

We see that the energy of the level is exponentially small compared with the depth of the well.

CHAPTER VII

THE QUASI-CLASSICAL CASE

CHAPTER VIII

SPIN

§ 54. Spin

IN BOTH classical and quantum mechanics, the law of conservation of angular momentum is a consequence of the isotropy of space with respect to a closed system. This already demonstrates the relation between the angular momentum and the symmetry properties under rotation. In quantum mechanics, however, the relation in question is a particularly far-reaching one, and essentially constitutes the basic content of the concept of angular momentum, especially as the classical definition of the angular momentum of a particle as the product $\boldsymbol{r} \times \boldsymbol{p}$ has no direct significance in quantum mechanics, owing to the fact that position and momentum cannot be simultaneously measured.

We have seen in §28 that, if the values of l and m are specified, the angular dependence of the wave function of the particle is determined, and therefore so are all its symmetry properties under rotation. The most general formulation of these properties involves specifying the transformation of the wave functions when the coordinate system is rotated.

The wave function ψ_{LM} of a system of particles (with specified values of the angular momentum L and its component M) remains unchanged[1] only in a rotation of the coordinate system about the z-axis. Any rotation that alters the direction of this axis has the result that the z-component of the angular momentum does not have a definite value. This means that, in the new coordinates, the wave function in general becomes a superposition (a linear combination) of $2L + 1$ functions corresponding to the different possible values of M for the given L. We can say that the $2L + 1$ functions ψ_{LM} are transformed into linear combinations of one another when the coordinate system is rotated.[2] The law governing this transformation (i.e.

---

[1] Apart from an unimportant phase factor.

[2] In mathematical terms, these functions are the irreducible representations of the rotation group. The number of functions which are transformed into linear combinations

the coefficients in the superposition as functions of the angles of rotation of the coordinate axes) is entirely determined by specifying the value of L. Thus the angular momentum acquires the significance of a quantum number which classifies the states of the system according to their transformation properties under rotation of the coordinate system. This aspect of the concept of angular momentum in quantum mechanics is particularly important because it is not directly related to the explicit angular dependence of the wave functions; the law of mutual transformation of these functions can be stated without reference to that dependence.

Let us consider a composite particle, such as an atomic nucleus, which is at rest as a whole and is in a definite internal state. In addition to an internal energy, it has also an angular momentum of definite magnitude L, due to the motion of the particles within the nucleus. This angular momentum can have $2L + 1$ different orientations in space. Thus, in considering the movement of a complex particle as a whole, we must assign to it, as well as its coordinates, another discrete variable: the projection of its internal angular momentum on some chosen direction in space.

However, with the preceding understanding of the concept of angular momentum, the origin of it becomes unimportant, and we naturally arrive at the concept of an "intrinsic" angular momentum which must be ascribed to the particle regardless of whether it is "composite" or "elementary".

Thus, in quantum mechanics an elementary particle must be assigned a certain "intrinsic" angular momentum unconnected with its motion in space. This property of elementary particles is peculiar to quantum theory (it disappears in the limit $\hbar \to 0$), and therefore has in principle no classical interpretation.[3]

The intrinsic angular momentum of a particle is called its *spin*, as distinct from the angular momentum due to the motion of the particle in space, called the *orbital angular momentum*.[4] The particle concerned may be either elementary, or composite but behaving in some respect as an elementary particle (e.g. an atomic nucleus). The spin of a particle (measured, like the orbital angular momentum, in units of $\hbar$) will be denoted by s.

For particles having spin, the description of the state by means of the wave function must determine the probability not only of its different positions in

---

of one another is called the dimension of the representation; it is assumed that this number cannot be made smaller by taking any other linear combinations of these functions.

[3] In particular, it would be wholly meaningless to imagine the "intrinsic" angular momentum of an elementary particle as being the result of its rotation "about its own axis".

[4] The physical idea that an electron has an intrinsic angular momentum was put forward by G. Uhlenbeck and S. Goudsmit in 1925. Spin was introduced into quantum mechanics in 1927 by W. Pauli.

space but also of the possible orientations of the spin. Thus the wave function must depend not only on three continuous variables, the coordinates of the particle, but also on a discrete spin variable, which gives the value of the projection of the spin on a selected direction in space (the z-axis) and takes a limited number of discrete values, which we shall denote by σ.

Let $\psi(x, y, z; \sigma)$ be such a wave function. It is essentially a set of several different functions of the coordinates, corresponding to different values of σ; these functions will be called the *spin components* of the wave function. The integral

$$\int |\psi(x, y, z; \sigma)|^2 \mathrm{d}V$$

determines the probability that the particle has a certain value of σ. The probability that the particle is in the volume element $\mathrm{d}V$ with any value of σ is

$$\mathrm{d}V \sum_\sigma |\psi(x, y, z; \sigma)|^2.$$

The quantum-mechanical spin operator, on being applied to the wave function, acts on the spin variable σ. In other words, it in some way linearly transforms the components of the wave function into one another. The form of this operator will be established later. However, it is easy to see from very general considerations that the operators $\widehat{s}_x, \widehat{s}_y, \widehat{s}_z$ satisfy the same commutation conditions as the operators of the orbital angular momentum.

The angular momentum operator is essentially the same as that of an infinitely small rotation. In deriving, in §26, the expression for the orbital angular momentum operator, we considered the result of applying the rotation operator to a function of the coordinates. In the case of the spin, this derivation becomes invalid, since the spin operator acts on the spin variable, and not on the coordinates. Hence, to obtain the required commutation relations, we must consider the operation of an infinitely small rotation in a general form, as a rotation of the system of coordinates. If we successively perform infinitely small rotations about the x-axis and the y-axis, and then about the same axes in the reverse order, it is easy to see by direct calculation that the difference between the results of these two operations is equivalent to an infinitely small rotation about the z-axis (through an angle equal to the product of the angles of rotation about the x and y-axes). We shall not pause here to carry out these simple calculations, as a result of which we again obtain the usual commutation relations between the operators of the components of angular momentum; these must therefore hold for the spin operators also:

$$\{\widehat{s}_y, \widehat{s}_z\} = \mathrm{i}\widehat{s}_x, \quad \{\widehat{s}_z, \widehat{s}_z\} = \mathrm{i}\widehat{s}_y, \quad \{\widehat{s}_x, \widehat{s}_y\} = \mathrm{i}\widehat{s}_z \qquad (54.1)$$

together with all the physical consequences resulting from them.

The commutation relations (54.1) enable us to determine the possible values of the absolute magnitude and components of the spin. All the results derived in §27 (formulae (27.7)–(27.9)) were based only on the commutation relations, and hence are fully applicable here also; we need only replace $\boldsymbol{L}$ in these formulae by $\boldsymbol{s}$. It follows from formula (27.7) that the eigenvalues of the component of the spin form a sequence of numbers differing by unity. However, we cannot now assert that these values must be integral, as we could for the component L_z of the orbital angular momentum (the derivation given at the beginning of §27 is invalid here, since it was based on the expression (26.14) for the operator $\widehat{l}_z$, which holds only for the orbital angular momentum).

Moreover, we find that the sequence of eigenvalues s_z is limited above and below by values equal in absolute magnitude and opposite in sign, which we denote by $\pm s$. The difference $2s$ between the greatest and least values of s_z must be an integer or zero. Consequently s can take the values $0, 1/2, 1, 3/2, \ldots$

Thus the eigenvalues of the square of the spin are

$$\boldsymbol{s}^2 = s(s+1) \tag{54.2}$$

where s can be either an integer (including zero) or half an integer. For given s, the component s_z of the spin can take the values $s, s-1, \ldots, -s$, i.e. $2s+1$ values in all. Accordingly, the wave function of a particle with spin s has $2s+1$ components.[5]

Experiment shows that the majority of the elementary particles (electrons, positrons, protons, neutrons, μ-mesons and all hyperons (Λ, Σ, Ξ)) have a spin of $1/2$. There are also elementary particles, the π-mesons and the K-mesons, whose spin is zero.

The total angular momentum of a particle is composed of its orbital angular momentum $\boldsymbol{l}$ and its spin $\boldsymbol{s}$. Their operators act on functions of different variables, and therefore, of course, commute. The eigenvalues of the total angular momentum

$$\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s} \tag{54.3}$$

are determined by the same "vector model" rule as the sum of the orbital angular momenta of two different particles (§31). That is, for given values of l

---

[5] Since s is fixed for each kind of particle, the spin angular momentum $\hbar s$ becomes zero in the limit of classical mechanics ($\hbar \to 0$). This consideration does not apply to the orbital angular momentum, since l can take any value. The transition to classical mechanics is represented by $\hbar$ tending to zero and l simultaneously tending to infinity, in such a way that the product $\hbar l$ remains finite.

and s, the total angular momentum can take the values $l+s, l+s-1, \ldots, |ls|$. Thus, for an electron (spin $1/2$) with non-zero orbital angular momentum l, the total angular momentum can be $j = l \pm 1/2$; for $l = 0$ the angular momentum j has, of course, only the one value $j = 1/2$.

The operator of the total angular momentum $\boldsymbol{J}$ of a system of particles is equal to the sum of the operators of the angular momentum $\boldsymbol{j}$ of each particle, so that its values are again determined by the vector model rules. The angular momentum $\boldsymbol{J}$ can be put in the form

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}, \quad \boldsymbol{L} = \sum_a \boldsymbol{l}_a, \quad \boldsymbol{S} = \sum_a \boldsymbol{s}_a, \tag{54.4}$$

where $\boldsymbol{S}$ may be called the total spin and $\boldsymbol{L}$ the total orbital angular momentum of the system. We notice that, if the total spin of the system is half-integral (or integral), the same is true of the total angular momentum, since the orbital angular momentum is always integral. In particular, if the system consists of an even number of similar particles, its total spin is always integral, and therefore so is the total angular momentum.

The operators of the total angular momentum $\boldsymbol{j}$ of a particle (or $\boldsymbol{J}$, of a system of particles) satisfy the same commutation rules as the operators of the orbital angular momentum or the spin, since these rules are general commutation rules holding for any angular momentum. The formulae (27.13) for the matrix elements of angular momentum, which follow from the commutation rules, are also valid for any angular momentum, provided that the matrix elements are defined with respect to the eigenstates of this angular momentum. Formulae (29.7)—(29.10) for the matrix elements of arbitrary vector quantities also remain valid (with appropriate change of notation).

PROBLEMS

A particle with spin $1/2$ is in a state with a definite value $s_z = 1/2$. Determine the probabilities of the possible values of the component of the spin along an axis z' at an angle θ to the z-axis.

SOLUTION. The mean spin vector $\overline{\boldsymbol{s}}$ is evidently along the z-axis and has magnitude $1/2$. Taking the component along the z'-axis, we find that the mean value of the spin in that direction is $\overline{s_{z'}} = (1/2)\cos\theta$. We also have $\overline{s_{z'}} = (1/2)(w_+ + w_-)$ where $w_\pm$ are the probabilities of the values $s_{z'} = \pm 1/2$. Since $w_+ + w_- = 1$, we find

$$w_+ = \cos 2\theta, w_- = \sin 2\theta.$$

§ 55. The spin operator

In the rest of this chapter we shall not be interested in the dependence of the wave functions on the coordinates. For example, in speaking of the

behaviour of the functions $\psi(x, y, z; \sigma)$ when the system of coordinates is rotated, we can suppose that the particle is at the origin, so that its coordinates remain unchanged by such a rotation, and the results obtained will characterize the behaviour of the function ψ with regard to the spin variable σ.

The variable σ differs from the ordinary variables (the coordinates) by being discrete. The most general form of a linear operator acting on functions of a discrete variable σ is

$$\left(\widehat{f}\psi\right)(\sigma) = \sum_{\sigma'} f_{\sigma\sigma'}\psi(\sigma') \tag{55.1}$$

where the $f_{\sigma\sigma'}$, are constants. We put ψ in parentheses in order to emphasize that the spin argument following it is not that of the original function ψ but that of the function resulting from it under the action of the operator $\widehat{f}$. It is easy to see that the quantities $f_{\sigma\sigma'}$, are the same as the matrix elements of the operator, defined by the usual rule (11.5).[6])

The integration over the coordinates in (11.5) is here replaced by summation over the discrete variable, so that the definition of the matrix element is

$$f_{\sigma_2\sigma_1} = \sum_{\sigma} \psi_{\sigma_2}^*(\sigma)\left[\widehat{f}\psi_{\sigma_1}(\sigma)\right] \tag{55.2}$$

Here $\psi_{\sigma_1}(\sigma)$ and $\psi_{\sigma_2}(\sigma)$ are the eigenfunctions of the operator $\widehat{s}_z$ corresponding to the eigenvalues $s_z = \sigma_1$ and σ_2; each such function corresponds to a state in which the particle has a definite value of s_z, i.e. in which only one component of the wave function is non-zero:[7])

$$\psi_{\sigma_1}(\sigma) = \delta_{\sigma\sigma_1}, \quad \psi_{\sigma_2}(\sigma) = \delta_{\sigma\sigma_2}. \tag{55.3}$$

According to (55.1),

$$\left(\widehat{f}\psi\right)(\sigma) = \sum_{\sigma'} f_{\sigma\sigma'}\psi_{\sigma_1}(\sigma') = \sum_{\sigma'} f_{\sigma\sigma'}\delta_{\sigma'\sigma_1} = f_{\sigma\sigma'}$$

and on substitution of this and $\psi_{\sigma_2}(\sigma)$ the equation (55.2) is satisfied identically; this completes the proof.

---

[6]) Note that the suffixes in the matrix elements on the right of (55.1) are written in an order which is, in a sense, the reverse of the usual order in (11.11).

[7]) More precisely, we should write

$$\psi_{\sigma_1}(\sigma) = \psi(x, y, z)\delta_{\sigma_1\sigma};$$

in (55.3) the coordinate factors are omitted, being unimportant in this connection.

We must once again emphasize the distinction between the specified eigenvalue σ_1 or σ_2 of s_z and the independent variable σ.

Thus the operators acting on functions of σ can be represented in the form of $(2s + 1)$-rowed matrices. In particular, we have for the operator of the spin itself, acting on the wave function, by (55.1),

$$(\widehat{\boldsymbol{s}}\psi)(\sigma) = \sum_{\sigma'} \boldsymbol{s}_{\sigma\sigma'}\psi(\sigma') \tag{55.4}$$

According to what has been said at the end of §54, the matrices $\widehat{s}_x, \widehat{s}_y, \widehat{s}_z$ are identical with the matrices $\widehat{L}_x, \widehat{L}_y, \widehat{L}_z$ obtained in §27, where the letters L and M need only be replaced by s and σ:

$$(s_x)_{\sigma,\sigma-1} = (s_x)_{\sigma-1,\sigma} = \frac{1}{2}\sqrt{(s+\sigma)(s-\sigma+1)},$$
$$(s_y)_{\sigma,\sigma-1} = -(s_y)_{\sigma-1,\sigma} = -\frac{i}{2}\sqrt{(s+\sigma)(s-\sigma+1)}, \tag{55.5}$$
$$(s_z)_{\sigma\sigma} = \sigma.$$

This determines the spin operator.

In the important case of a spin of $1/2$ ($s = 1/2, = \pm 1/2$), these matrices have two rows, and are of the form

$$\widehat{\boldsymbol{s}} = \frac{1}{2}\widehat{\boldsymbol{\sigma}} \tag{55.6}$$

where[8]

$$\widehat{\boldsymbol{\sigma}}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \widehat{\boldsymbol{\sigma}}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \widehat{\boldsymbol{\sigma}}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{55.7}$$

These are called *Pauli matrices*. The matrix $\widehat{\boldsymbol{s}}_z = \frac{\widehat{\sigma}_z}{2}$ is diagonal, as it should be, since it is defined in terms of the eigenfunctions of the quantity $\boldsymbol{s}_z$ itself.[9]

---

[8] In the tabular matrices (55.7) the rows and columns are numbered by the values of , the row number corresponding to the first and the column number to the second suffix of the matrix element. In the present case, these numbers are $+1/2$ and $-1/2$. The action of the. operator shown by (55.4) multiplies row σ of the matrix by a column matrix containing the components of the wave function:

$$\psi = \begin{pmatrix} \psi(1/2) \\ \psi(-1/2) \end{pmatrix}$$

[9] There should be no misunderstanding because of the use of the same letter to denote the spin component and the Pauli matrices, since the latter always have the circumflex.

The following are some specific properties of the Pauli matrices. Direct multiplication of the matrices (55.7) gives the equations

$$\widehat{\boldsymbol{\sigma}}_x^2 = \widehat{\boldsymbol{\sigma}}_y^2 = \widehat{\boldsymbol{\sigma}}_z^2 = 1$$
$$\widehat{\boldsymbol{\sigma}}_y\widehat{\boldsymbol{\sigma}}_z = \mathrm{i}\widehat{\boldsymbol{\sigma}}_x, \quad \widehat{\boldsymbol{\sigma}}_z\widehat{\boldsymbol{\sigma}}_x = \mathrm{i}\widehat{\boldsymbol{\sigma}}_y, \quad \widehat{\boldsymbol{\sigma}}_x\widehat{\boldsymbol{\sigma}}_y = \mathrm{i}\widehat{\boldsymbol{\sigma}}_z \tag{55.8}$$

Combining these with the general commutation rules (54.1), we find that

$$\widehat{\boldsymbol{\sigma}}_\mathrm{i}\widehat{\boldsymbol{\sigma}}_k + \widehat{\boldsymbol{\sigma}}_k\widehat{\boldsymbol{\sigma}}_\mathrm{i} = 2\delta_{\mathrm{i}k} \tag{55.9}$$

i.e. the Pauli matrices anticommute with one another. By means of these equations, we can easily verify the following useful formulae:

$$\widehat{\boldsymbol{\sigma}}^2 = 3, \quad (\widehat{\boldsymbol{\sigma}} \cdot \boldsymbol{a})(\widehat{\boldsymbol{\sigma}} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot \boldsymbol{b} + \mathrm{i}\widehat{\boldsymbol{\sigma}} \cdot \boldsymbol{a} \times \boldsymbol{b}. \tag{55.10}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are any vectors.[10]) According to these relations, any scalar polynomial formed from the matrices $\widehat{\boldsymbol{\sigma}}_i$ can be reduced to terms independent of $\widehat{\boldsymbol{\sigma}}$ and terms linear in $\widehat{\boldsymbol{\sigma}}$; hence it follows that any scalar function of the operator reduces to a linear function (see Problem 1). Lastly, the values of the traces (sums of diagonal elements) of the Pauli matrices and their products are

$$\mathrm{tr}\,\widehat{\sigma}_i = 0, \quad \mathrm{tr}\,\widehat{\sigma}_i\widehat{\sigma}_k = 2\delta_{ik}. \tag{55.11}$$

Subsequent sections of this chapter give a more detailed account of the spin properties of wave functions, including their behaviour under any rotation of the coordinate system, but we may note immediately an important property of these functions, namely their behaviour in respect of rotations about the z-axis.

Let there be an infinitesimal rotation through an angle $\delta\varphi$ about the z-axis. The operator of such a rotation is expressed in terms of the angular momentum operator (in this case, the spin operator) as $1 + \mathrm{i}\delta\varphi \cdot \widehat{s}_z$. As a result of the rotation, the functions $\psi(\sigma)$ therefore become $\psi(\sigma) + \delta\psi(\sigma)$, where

$$\delta\psi(\sigma) = \mathrm{i}\delta\varphi \cdot \widehat{s}_z\psi(\sigma) = \mathrm{i}\sigma\psi(\sigma)\delta\varphi$$

Writing this relation in the form $\mathrm{d}\psi/\mathrm{d}\varphi = \mathrm{i}\sigma\psi(\sigma)$ and integrating, we find that a rotation through a finite angle φ changes the functions $\psi(\sigma)$ into

$$\psi(\sigma)' = \psi(\sigma)\mathrm{e}^{\mathrm{i}\sigma\varphi} \tag{55.12}$$

In particular, a rotation through 2π multiplies them by a factor $\mathrm{e}^{2\pi\mathrm{i}\sigma}$, which is the same for all σ and is equal to $(-1)^{2s}$ (since 2σ always has the same

---

[10]) The terms on the right of (55.8)–(55.10) which are independent of $\widehat{\boldsymbol{\sigma}}$ must, of course, be understood as constants multiplying the unit two-by-two matrix.

parity as $2s$). Thus, in a complete rotation of the coordinate system about the z-axis, the wave functions of particles with integral spin return to their original values, and those of particles with half-integral spin change sign.

PROBLEMS

1. Reduce an arbitrary function of the scalar $a + \boldsymbol{b} \cdot \widehat{\boldsymbol{\sigma}}$ linear in the Pauli matrices to another linear function.

SOLUTION. TO determine the coefficients in the required formula

$$f(a + \boldsymbol{b} \cdot \widehat{\boldsymbol{\sigma}}) = \boldsymbol{A} + \boldsymbol{B}\widehat{\boldsymbol{\sigma}},$$

we note that, when the z-axis is taken in the direction of $\boldsymbol{b}$, the eigenvalues of the operator $a + \boldsymbol{b} \cdot \widehat{\boldsymbol{\sigma}}$ are $a \pm b$, and the corresponding eigenvalues of the operator $f(a + \boldsymbol{b} \cdot \widehat{\boldsymbol{\sigma}})$ are $f(a \pm b)$. Hence we find

$$A = \frac{1}{2}[f(a + b) + f(ab)], \boldsymbol{B} = \frac{\boldsymbol{b}}{2b}[f(a + b) - f(ab)].$$

2. Determine the values of the scalar product $\boldsymbol{S}_1 \cdot \boldsymbol{S}_2$ of spins $(1/2)$ of two particles in states in which the total spin of the system, $\boldsymbol{S} = \boldsymbol{s}_1 + \boldsymbol{s}_2$, has definite values (0 or 1).

SOLUTION. From the general formula (31.3), which is valid for the addition of any two angular momenta, we find

$$\boldsymbol{s}_1 \cdot \boldsymbol{s}_2 = 1/4 \text{ for } S = 1, \quad \boldsymbol{s}_1 \cdot \boldsymbol{s}_2 = -3/4 \text{ for } S = 0.$$

3. Which powers of the operator $\hat{s}$ of an arbitrary spin $\widehat{\boldsymbol{s}}$ are independent?

SOLUTION. The operator

$$(\widehat{s}_z - s)(\widehat{s}_z - s + 1) \ldots (\widehat{s}_z + s),$$

formed from the differences between $\widehat{s}_z$ and all possible eigenvalues s_z, gives zero when it is applied to any wave function, and is therefore itself zero. Hence it follows that $(\widehat{s}_z)^{2s+1}$ is expressed in terms of lower powers of the operator $\widehat{s}_z$, so that only its powers from 1 to $2s$ are independent.

§ 56. Spinors

When the spin is zero, the wave function has only one component, $\psi(0)$. The effect of the spin operator is to reduce it to zero: $\widehat{\boldsymbol{s}}\psi = 0$. The relation between $\widehat{\boldsymbol{s}}$ and the operator of an infinitesimal rotation implies that the wave function of a particle with zero spin is invariant under rotation of the coordinate system, i.e. it is a scalar.

The wave function of a particle with spin $1/2$ has two components, $\psi(1/2)$ and $\psi(-1/2)$. For convenience in later generalizations, we shall distinguish these components by the superscripts 1 and 2 respectively. The two-component quantity

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} \equiv \begin{pmatrix} \psi\left(1/2\right) \\ \psi\left(-1/2\right) \end{pmatrix} \tag{56.1}$$

is called a *spinor*.

In any rotation of the coordinate system, the components of the spinor undergo a linear transformation:

$$\psi^{1'} = a\psi^1 + b\psi^2, \quad \psi^{2'} = c\psi^1 + d\psi^2. \tag{56.2}$$

This may be written

$$\psi^{\lambda'} = (\widehat{U}\psi)^\lambda, \quad \widehat{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{56.3}$$

where $\widehat{U}$ is the transformation matrix.[11] Its elements are in general complex functions of the angles of rotation of the coordinate axes. They are connected by relations which follow directly from the physical conditions imposed on the spinor as the wave function of a particle.

Let us consider the bilinear form

$$\psi^1\varphi^2 - \psi^2\varphi^1, \tag{56.4}$$

where ψ and φ are two spinors. A simple calculation gives

$$\psi^{1'}\varphi^{2'} - \psi^{2'}\varphi^{1'} = (ad - bc)(\psi^1\varphi^2 - \psi^2\varphi^1),$$

i.e. (56.4) is transformed into itself when the coordinate system is rotated. If, however, there is only one function which is transformed into itself, it can be regarded as corresponding to zero spin, and therefore must be a scalar, i.e. must remain unchanged when the coordinate system is rotated in any manner. Hence we have

$$ad - bc = 1; \tag{56.5}$$

the determinant of the transformation matrix is unity.[12]

---

[11]) The notation $\widehat{U}\psi$ implies that the rows of the matrix $\widehat{U}$ are multiplied by the column ψ.

[12]) Such a transformation of two quantities is called a *binary transformation*.

Further relations follow from the requirement that the expression

$$\psi^1 \psi^{1*} + \psi^2 \psi^{2*}, \tag{56.6}$$

which determines the probability of finding the particle at a given point in space, should be a scalar. A transformation which leaves unchanged the sum of the squared moduli of the quantities is a unitary transformation, i.e. we must have $\widehat{U}^\dagger = \widehat{U}^{-1}$ (see §12). With the condition (56.5) the inverse matrix is

$$\widehat{U}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Equating this to the Hermitian conjugate matrix

$$\widehat{U}^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix},$$

we find

$$a = d^*, \quad b = -c^* \tag{56.7}$$

By virtue of the relations (56.5) and (56.7), the four complex quantities a, b, c, d actually contain only three independent real parameters, corresponding to the three angles which define a rotation of a three-dimensional system of coordinates.

Comparison of the expressions for the scalars (56.4) and (56.6) shows that ψ^{1*} and ψ^{2*} must be transformed as ψ^2 and $-\psi^1$ respectively. It is easy to verify that this is in fact so, using (56.5) and (56.7).[13]

It is possible to put the algebra of spinors in a form analogous to that of tensor algebra. This is done by introducing, in addition to *contravariant* spinor components ψ^1, ψ^2 (with superscripts), the *covariant* components (with subscripts), defined by

$$\psi_1 = \psi^2, \quad \psi_2 = -\psi^1 \tag{56.8}$$

The invariant combination (56.4) of the two spinors may also be written as a scalar product

$$\psi^\lambda \varphi_\lambda = \psi^1 \varphi_1 + \psi^1 \varphi_2 - \psi^2 \varphi_1; \tag{56.9}$$

---

[13] This property is closely associated with symmetry under time reversal. The latter (see §18) involves the replacement of the wave function by its complex conjugate. Under time reversal, the angular momentum components also change sign. Hence the functions that are the complex conjugates of the components $\psi^1 \equiv \psi(1/2)$ and $\psi^2 \equiv \psi(-1/2)$ must have properties equivalent to those of the components corresponding to spin projections $-1/2$ and $1/2$ respectively.

here and below, summation over repeated (*dummy*) indices is implied, as in tensor algebra. We may note the following rule which has to be borne in mind in spinor algebra. We have

$$\psi^\lambda \varphi_\lambda = \psi^1 \varphi_1 + \psi^2 \varphi_2 = -\psi_2 \varphi^2 - \psi_1 \varphi^1.$$

Thus

$$\psi^\lambda \varphi_\lambda = -\psi_\lambda \varphi^\lambda \qquad (56.10)$$

Hence it is evident that the scalar product of any spinor with itself is zero:

$$\psi^\lambda \psi_\lambda = 0. \qquad (56.11)$$

According to the foregoing discussion, the quantities ψ_1 and ψ_2 are transformed as ψ^{1*} and ψ^{2*}, i.e.

$$\psi'_\lambda = \left(\widehat{U}^* \psi \right)_\lambda \qquad (56.12)$$

The product $\widehat{U}^* \psi$ may also be written as $\psi \widetilde{\widehat{U}^*}$, with the transposed matrix $\widetilde{\widehat{U}^*}$. Since $\widehat{U}$ is unitary, we have $\widetilde{\widehat{U}^*} = \widehat{U}^{-1}$, so that $\psi'_\lambda = \left(\psi \widehat{U}^{-1} \right)_\lambda$ or[14]

$$\psi_\lambda = \left(\psi' \widehat{U} \right)_\lambda \qquad (56.13)$$

Analogously to the transition from vectors to tensors in ordinary tensor algebra, we can introduce the idea of spinors of higher rank. Thus, a quantity $\psi^{\lambda\mu}$, having four components which are transformed as the products $\psi^\lambda \varphi^\mu$ of the components of two spinors of rank one, is called a spinor of rank two. Besides the contra variant components $\psi^{\lambda\mu}$ we can consider the covariant components $\psi^{\lambda\mu}$ and the mixed components ψ^μ_λ which are transformed as the products $\psi_\lambda \varphi_m u$ and $\psi_\lambda \varphi^\mu$ respectively. Spinors of any rank are similarly defined.

The transition from contravariant to covariant spinor components and *vice versa* may be written

$$\psi_\lambda = g_{lambd\mu} \psi^\mu, \quad \psi^\lambda = g^{\mu\lambda} \psi_m u, \qquad (56.14)$$

where

$$(g_{\lambda\mu}) = \left(g^{\lambda\mu} \right) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad (56.15)$$

---

[14] The notation $\psi \widehat{U}$ (with ψ to the left of $\widehat{U}$) denotes that the components $(\psi_1 \psi_2)$ as a row are multiplied by the columns of the matrix $\widehat{U}$.

is the *metric spinor* in a vector space of two dimensions. Thus we have, for example,

$$\psi^\mu_\lambda = g_{\lambda\nu}\psi^{\nu\mu}, \quad \psi_{\lambda\mu} = g_{\lambda\nu}g_{\mu\rho}\psi^{\nu\rho}$$

so that $\psi_{12} = -\psi_1{}^1 = -\psi^{21}$, $\psi_{11} = \psi_1{}^2 = \psi^{22}$, and so on.

The quantities $g_{\lambda\mu}$ themselves form an antisymmetric unit spinor of rank two. It is easy to see that the values of its components remain unchanged under transformations of the coordinates, and that

$$g_{\lambda\nu}g^{\mu\nu} = \delta^\mu_\lambda \tag{56.16}$$

where $\delta_1^1 = \delta_2^2 = 1, \delta_2^1 = \delta_1^2 = 0$.

As in ordinary tensor algebra, there are two fundamental operations in spinor algebra: multiplication, and contraction with respect to a pair of indices. The *multiplication* of two spinors gives a spinor of higher rank; thus, from two spinors of ranks two and three, $\psi_{\lambda\mu}$ and $\psi^{\nu\rho\sigma}$, we can form a spinor of rank five, $\psi_{\lambda\mu}\psi^{\nu\rho\sigma}$. *Contraction* with respect to a pair of indices (i.e. summation of the components over corresponding values of one covariant and one contravariant index) decreases the rank of a spinor by two. Thus, a contraction of the spinor $\psi_{\lambda\mu}{}^{\nu\rho\sigma}$ with respect to the indices μ and ν gives the spinor $\psi_{\lambda\mu}{}^{\mu\rho\sigma}$ of rank three; the contraction of the spinor $\psi_\lambda{}^\mu$ gives the scalar $\psi_\lambda{}^\lambda$. Here there is a rule similar to that expressed by formula (56.10): if we interchange the upper and lower indices with respect to which the contraction is effected, the sign is changed (i.e. $\psi_\lambda{}^\lambda = -\psi^\lambda{}_\lambda$). Hence, in particular, it follows that, if a spinor is symmetrical with respect to any two of its indices, the result of a contraction with respect to these indices is zero. Thus, for a symmetrical spinor $\psi_{\lambda\mu}$ of rank two, we have $\psi_\lambda{}^\lambda = 0$.

A spinor of rank n symmetrical with respect to all its indices is called a *symmetrical spinor* of rank n. From an asymmetrical spinor we can construct a symmetrical one by the process of *symmetrization*, i.e. summation of the components obtained by all possible interchanges of the indices. From what has been said above, it is impossible to construct (by contraction) a spinor of lower rank from the components of a symmetrical spinor.

Only a spinor of rank two can be antisymmetrical with respect to all its indices. For, since each index can take only two values, at least two out of three or more indices must have the same value, and therefore the components of the spinor are zero identically. Any antisymmetrical spinor of rank two is a scalar multiple of the unit spinor $g_{\lambda\mu}$. We may notice here the following relation:

$$g_{\lambda\mu}\psi_\nu + g_{\mu\nu}\psi_\lambda + g_{\nu\lambda}\psi_\mu = 0 \tag{56.17}$$

(where ψ_λ is any spinor), which follows from the above; this rule is simply a

consequence of the fact that the expression on the left is (as we may easily verify) an antisymmetrical spinor of rank three.

The spinor which is the product of a spinor $\psi_{\lambda\mu}$ with itself, on contraction with respect to one pair of indices, becomes antisymmetrical with respect to the other pair:

$$\psi_{\lambda\nu}\psi_\mu{}^\nu = -\psi_\lambda{}^\nu\psi_{\mu\nu}.$$

Hence, from what was said above, this spinor must be a scalar multiple of the spinor $g_{\lambda\mu}$. Defining the scalar factor so that contraction with respect to the second pair of indices gives the correct result, we find

$$\psi_{\lambda\nu}\psi_n u^\nu = -(1/2)\psi_{\rho\sigma}\psi^{\rho\sigma}g_{\lambda\mu} \tag{56.18}$$

The components of the spinor $\psi_{\lambda\mu\ldots}{}^*$ which is the complex conjugate of $\psi_{\lambda\mu\ldots}$ are transformed as the components of a contravariant spinor $\varphi^{\lambda\mu\cdots}$, and conversely. The sum of the squared moduli of the components of any spinor is consequently invariant.

§ 57. The wave functions of particles with arbitrary spin

Having developed a formal algebra for spinors of any rank, we can now turn to our immediate problem, to study the properties of wave functions of particles with arbitrary spin.

This subject is conveniently approached by considering an assembly of n particles with spin 1/2. The greatest possible value of the z-component of the total spin is $n/2$, which is obtained when $s_z = 1/2$ for every particle (i.e. all the spins are directed the same way, along the z-axis). In this case we can evidently say that the total spin S of the system is also $n/2$.

All the components of the wave function $\psi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ of the system of particles are then zero, except for $\psi(1/2, 1/2, \ldots, 1/2)$. If we write the wave function as a product of n spinors $\psi^\lambda\psi^\mu \ldots$, each of which refers to one of the particles, only the component with $\lambda, \mu, \cdots = 1$ in each spinor is not zero. Thus only the product $\psi^1\varphi^1$ is not zero. The set of all these products, however, is a spinor of rank n which is symmetrical with respect to all its indices. If we transform the coordinate system (so that the spins are not directed along the z-axis), we obtain a spinor of rank n, general in form except that it is symmetrical as before.

The spin properties of wave functions, being essentially their properties with respect to rotations of the coordinate system, are identical for a particle with spin s and for a system of $n = 2s$ particles each with spin 1/2 directed

so that the total spin of the system is s. Hence we conclude that the wave function of a particle with spin s is a symmetrical spinor of rank $n = 2s$.

It is easy to see that the number of independent components of a symmetrical spinor of rank $2s$ is equal to $2s + 1$, as it should be. For all those components are the same whose indices include $2s$ ones and 0 twos; so are all those with $2s1$ ones and 1 two, and so on up to 0 ones and $2s$ twos.

Mathematically, the symmetrical spinors provide a classification of the possible types of transformation of quantities when the coordinate system is rotated. If there are $2s+1$ different quantities which are transformed linearly into one another (and which cannot be reduced in number by any choice of linear combinations of them), then we can assert that their law of transformation is equivalent to that of the components of a symmetrical spinor of rank $2s$. Any set of any number of functions which are transformed linearly into one another when the coordinate system is rotated can be reduced (by an appropriate linear transformation) to one or more symmetrical spinors.[15])

Thus an arbitrary spinor $\psi_{\lambda\mu\nu}$ of rank n can be reduced to symmetrical spinors of ranks $n, n-2, n-4, \ldots$. In practice, such a reduction can be made as follows. By symmetrizing the spinor $\psi_{\lambda\mu\nu}$ with respect to all its indices, we form a symmetrical spinor of the same rank n. Next, by contracting the original spinor $\psi_{\lambda\mu\nu}$ with respect to various pairs of indices, we obtain spinors of rank $n - 2$, of the form $\psi^{\lambda}{}_{\lambda\nu\ldots}$, which, in turn, we symmetrize, so that symmetrical spinors of rank $n - 2$ are obtained. By symmetrizing the spinors obtained by contracting $\psi_{\lambda\mu\ldots}$ with respect to two pairs of indices, we obtain symmetrical spinors of rank $n - 4$, and so on.

We have still to establish the relation between the components of a symmetrical spinor of rank $2s$ and the $2s + 1$ functions $\psi(\sigma)$, where $\sigma = s, s - 1, \ldots, -s$. The component

$$\psi_{\underbrace{11\ldots1}_{s+\sigma}\overbrace{22\ldots2}^{s+\sigma}},$$

in whose indices 1 occurs $s + \sigma$ times and 2 $s - \sigma$ times, corresponds to a value σ of the projection of the spin on the z-axis. For, if we again consider a system of $n = 2s$ particles with spin $1/2$, instead of one particle with spin s, the product corresponds to the above component

$$\overbrace{\psi^1\varphi^1}^{s+\sigma}\ldots\overbrace{\chi^2\rho^2}^{s-\sigma}$$

---

[15]) In other words, the symmetrical spinors form what are called irreducible representations of the rotation group (see §98).

this product belongs to a state in which $s+\sigma$ particles have a projection of the spin equal to $+1/2$, and $s-$ a projection of $-1/2$, so that the total projection is $1/2(s+\sigma)-1/2(s-\sigma) = \sigma$. Finally, the proportionality coefficient between the above component of the spinor and $\psi(\sigma)$ is chosen so that the equation

$$\sum_{\sigma=-s}^{+s} |\psi(\sigma)|^2 = \sum_{\lambda,\mu,\cdots=1}^{2} |\psi^{\lambda\mu\cdots}|^2 \qquad (57.1)$$

holds; this sum is a scalar, as it should be, since it determines the probability of finding the particle at a given point in space. In the sum on the right-hand side, the components with $(s+\sigma)$ indices 1 occur

$$\frac{(2s)!}{(s+\sigma)!(s-\sigma)!}$$

times. Hence it is clear that the relation between the functions $\psi(\sigma)$ and the components of the spinor is given by the formula

$$\psi(\sigma) = \sqrt{\frac{(2s)!}{(s+\sigma)!(s-\sigma)!}}\, \psi^{\overbrace{\underbrace{11\ldots1}_{s+\sigma}\overbrace{22\ldots2}^{s-\sigma}}} \qquad (57.2)$$

The relation (57.2) ensures the fulfilment not only of the condition (57.1), but also, as we easily see, of the more general condition

$$\psi^{\lambda\mu\cdots}\varphi_{\lambda\mu\cdots} = \sum_{\sigma}(-1)^{s-\sigma}\psi(\sigma)\varphi(-\sigma), \qquad (57.3)$$

where $\psi^{\lambda\mu\cdots}$ and $\varphi_{\lambda\mu\cdots}$ are two different spinors of the same rank, while $\psi(\sigma)$, $\varphi(\sigma)$ are functions derived from these spinors by formula (57.2); the factor $(-1)^{s-\sigma}$ is due to the fact that, when all the indices of the spinor components are raised, the sign changes as many times as there are twos among the indices.

Formulae (55.5) determine the result of the action of the spin operator on the wave functions $\psi(\sigma)$. It is not difficult to find how these operators act on a wave function written in the form of a spinor of rank $2s$. For a spin $1/2$, the functions $\psi(+1/2)$, $\psi(-1/2)$ are the same as the components ψ^1, ψ^2 of the spinor. According to (55.6) and (55.7), the result of the spin operators' acting on them will be

$$\begin{aligned}
(\widehat{s}_x\psi)^1 &= (1/2)\psi^2, & (\widehat{s}_y\psi)^1 &= -(\mathrm{i}/2)\psi^2, & (\widehat{s}_z\psi)^1 &= (1/2)\psi^1, \\
(\widehat{s}_x\psi)^2 &= (1/2)\psi^1, & (\widehat{s}_y\psi)^2 &= (\mathrm{i}/2)\psi^1, & (\widehat{s}_x\psi)^2 &= -(1/2)\psi^2.
\end{aligned} \qquad (57.4)$$

To pass to the general case of arbitrary spin, we again consider a system of $2s$ particles with spin $1/2$, and write its wave function as a product of $2s$ spinors. The spin operator of the system is the sum of the spin operators of each particle, acting only on the corresponding spinor, the result of this action being given by formulae (57.4). Next, returning to arbitrary symmetrical spinors, i.e. to the wave functions of a particle with spin s, we obtain

$$(\widehat{s}_x \psi)\overbrace{11}^{s+\sigma}\dots\overbrace{22}^{s-\sigma}\dots = \frac{s+\sigma}{2}\psi\overbrace{11}^{s+\sigma-1}\dots\overbrace{22}^{s-\sigma+1}\dots + \frac{s-\sigma}{2}\psi\overbrace{11}^{s+\sigma+1}\dots\overbrace{22}^{s-\sigma-1}\dots,$$

$$(\widehat{s}_y \psi)\overbrace{11}^{s+\sigma}\dots\overbrace{22}^{s-\sigma}\dots = -\mathrm{i}\frac{s+\sigma}{2}\psi\overbrace{11}^{s+\sigma-1}\dots\overbrace{22}^{s-\sigma+1}\dots + \mathrm{i}\frac{s-\sigma}{2}\psi\overbrace{11}^{s+\sigma+1}\dots\overbrace{22}^{s-\sigma-1}\dots, \quad (57.5)$$

$$(\widehat{s}_z \psi)\overbrace{11}^{s+\sigma}\dots\overbrace{22}^{s-\sigma}\dots = \sigma\psi\overbrace{11}^{s+\sigma}\dots\overbrace{22}^{s-\sigma}\dots.$$

Hitherto we have spoken of spinors as wave functions of the intrinsic angular momentum of elementary particles. Formally, however, there is no difference between the spin of a single particle and the total angular momentum of any system regarded as a whole, neglecting its internal structure. It is therefore evident that the transformation properties of spinors apply equally to the behaviour, with respect to rotations in space, of the wave functions ψ_{jm} of any particle or system of particles with total angular momentum j, independent of whether orbital or spin angular momentum is concerned. There must therefore be some definite relation between the laws of transformation for the eigenfunctions ψ_{jm} under rotations of the coordinate system and those for the components of a symmetrical spinor of rank $2j$.

In establishing this relation we must, however, make a clear distinction between two aspects of the dependence of the wave functions on the component m (for a given value of j). The wave function may be regarded as the probability amplitude for various values of m, or may be considered for a given value of m.

These two aspects have already been discussed at the beginning of §55, where we dealt with the eigenfunction $\delta_{\sigma\sigma_0}$ of the operator $\widehat{s}_z$ which corresponds to $s_z = \sigma_0$. The mathematical difference between them is especially clear for a particle of spin $s = 1/2$. In this case the spin function is, with respect to the variable σ, a contravariant spinor of rank 1, i.e. must be written in spinor notation as $\delta_{\sigma_0}^{\sigma}$. With respect to σ_0 it is therefore a covariant spinor.

This is evidently a general result: the eigenfunctions ψ_{jm} can be put in correspondence with the components of a covariant symmetrical spinor of

rank $2j$ by means of formulae analogous to (57.2):[16])

$$\psi_{jm} = \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}}\psi\underbrace{11\ldots}_{j+m}\underbrace{22\ldots}_{j-m}. \tag{57.6}$$

The eigenfunctions of integral angular momentum j are spherical harmonics Y_{jm}. The case $j = 1$ is of particular importance. The three spherical harmonics Y_{1m} are

$$Y_{10} = \mathrm{i}\sqrt{\frac{3}{4\pi}}\cos\theta = \mathrm{i}\sqrt{\frac{3}{4\pi}}n_z,$$

$$Y_{1,\pm 1} = \mp\mathrm{i}\sqrt{\frac{3}{8\pi}}\sin\theta \cdot \mathrm{e}^{\pm\mathrm{i}\varphi} = \mp\mathrm{i}\sqrt{\frac{3}{8\pi}}(n_x \pm \mathrm{i}n_y)$$

where $\boldsymbol{n}$ is a unit vector along the radius vector. It is seen that these three functions are equivalent, as regards their transformation properties, to the components of a vector $\boldsymbol{a}$, with the relations

$$\psi_{10} = \mathrm{i}a_z, \quad \psi_{11} = -\frac{\mathrm{i}}{\sqrt{2}}(a_x + \mathrm{i}a_y), \quad \psi_{1,-1} = \frac{\mathrm{i}}{\sqrt{2}}(a_x - \mathrm{i}a_y). \tag{57.7}$$

Comparing with (57.6), we see that the components of a symmetrical spinor of rank two can be brought into correspondence with the components of thevector by the formulae

$$\psi_{12} = \frac{\mathrm{i}}{\sqrt{2}}a_z, \quad \psi_{11} = -\frac{\mathrm{i}}{\sqrt{2}}(a_x + \mathrm{i}a_y), \quad \psi_{22} = \frac{\mathrm{i}}{\sqrt{2}}(a_x - \mathrm{i}a_y), \tag{57.8}$$

$$\psi^{12} = -\frac{\mathrm{i}}{\sqrt{2}}a_z, \quad \psi^{11} = \frac{\mathrm{i}}{\sqrt{2}}(a_x - \mathrm{i}a_y), \quad \psi^{22} = -\frac{\mathrm{i}}{\sqrt{2}}(a_x + \mathrm{i}a_y). \tag{57.9}$$

Conversely

$$a_z = \mathrm{i}\sqrt{2}\psi^{12}, \quad a_x = \frac{\mathrm{i}}{\sqrt{2}}(\psi^{22} - \psi^{11}), \quad a_y = \frac{1}{\sqrt{2}}(\psi^{11} + \psi^{22}). \tag{57.10}$$

---

[16]) This result can also be regarded somewhat differently. If the wave function ψ of a particle in a state with angular momentum j is expanded in terms of the eignfunctions ψ_{jm}:

$$\psi = \sum_m a_m\psi_{jm},$$

then the coefficients a_m are the probability amplitudes for various values of m. In this sense they correspond to the "components" $\psi(m)$ of a spin wave function, and this gives their law of transformation. On the other hand, the value of ψ at a given point in space cannot depend on the choice of the coordinate system, i.e. the sum $\sum a_m\psi_{jm}$ must be a scalar. Comparing with the scalar (57.3), we see that a_m must transform as $(-1)^{j-m}\psi_{j,-m}$.

It is easily verified that with these definitions we have

$$\psi_{\lambda\mu}\varphi^{\lambda\mu} = \boldsymbol{a}\boldsymbol{b} \tag{57.11}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors corresponding to the symmetrical spinors $\psi^{\lambda\mu}$ and $\varphi^{\lambda\mu}$. It is also not difficult to see that there is a correspondence between the spinor and the vector[17]

$$\psi_\nu^\lambda \varphi^{\mu\nu} + \psi_\nu^\mu \varphi^{\lambda\nu} \text{ and } \sqrt{2}\boldsymbol{a} \times \boldsymbol{b}. \tag{57.12}$$

Formulae (57.10) may be compactly written by means of the Pauli matrices:

$$\boldsymbol{a} = \frac{\mathrm{i}}{\sqrt{2}}\boldsymbol{\sigma}^\lambda{}_\mu\varphi^\mu_\lambda \quad \varphi^\mu_\lambda = -\frac{\mathrm{i}}{\sqrt{2}}\boldsymbol{a}\cdot\boldsymbol{\sigma}^\mu_\lambda \tag{57.13}$$

the matrix indices of $\widehat{\boldsymbol{\sigma}}$ are written as superscript and subscript in correspondence with the position of the spinor indices in ψ^μ_λ. The origin of this formula is easily understood by considering the particular case where the spinor of rank two ψ^μ_λ reduces to a product of a spinor of rank one ψ^μ and its complex conjugate $\psi^{\lambda*}$. Then the quantity

$$\frac{1}{2}\psi^{\lambda*}\boldsymbol{\sigma}^\mu{}_\lambda\psi^\nu$$

is the mean value of the spin (for a particle with wave function $\psi^m u$) and it is therefore evidently a vector.

The relations (57.8) or (57.9) are a particular case of a general rule: any symmetrical spinor of even rank $2j$, where j is integral, can be correlated with a symmetrical tensor of half the rank (j) which gives zero on contraction with respect to any pair of indices; we call this an *irreducible tensor*.

This follows from the fact that the numbers of independent components of the spinor and of the tensor are the same ($2j+1$), as may easily be seen.[18] The relation between the components of the spinor and of the tensor can be found by means of formulae (57.8)–(57.10), if we consider a spinor of the rank concerned as the product of several spinors of rank two, and the tensor as a product of vectors.

PROBLEMS

1. Rewrite the definition (57.4) of the operator of spin $1/2$ in terms of the spinor components of the vector $\widehat{s}$.

---

[17] The mixed components of a symmetrical spinor may be written in the form ψ^λ_μ, without distinction between $\psi^\lambda{}_\mu$ and $\psi_\mu{}^\lambda$.

[18] We can say that the $2j+1$ components of an irreducible tensor of rank j (an integer), the $2j+1$ spherical harmonics Yjm, and the $2j+1$ components of a symmetrical spinor of rank $2j$ give the same irreducible representation of the rotation group.

SOLUTION. By means of formulae (57.9), which give the relation between the vector $\widehat{s}^{\lambda\mu}$ and the spinor $\widehat{s}^{\lambda\mu}$, the definition (57.4) can be written as

$$\widehat{s}^{\lambda\mu}\psi^{\nu} = \frac{\mathrm{i}}{2\sqrt{2}}(\psi^{\lambda}g^{\mu\nu} + \psi^{\mu}g^{\lambda\nu})$$

2. Derive formulae which determine the effect of the spin operator on a vector wave function of a particle with spin 1.

SOLUTION. The relation between the components of the vector function $\boldsymbol{\psi}$ and the components of the spinor $\psi^{\lambda\mu}$ is given by formulae (57.9), and from (57.5) we have

$$\widehat{s}_z\psi_+ = -\psi_+, \quad \widehat{s}_z\psi_- = \psi_-, \quad \widehat{s}_z\psi_z = 0.$$

(where $\psi_{\pm} = \psi_x \pm \mathrm{i}\psi_y$) or

$$\widehat{s}_z\psi_x = -\mathrm{i}\psi_y, \quad \widehat{s}_z\psi_y = \mathrm{i}\psi_x, \quad \widehat{s}_z\psi_z = 0.$$

The remaining formulae are derived from these by cyclic permutation of the suffixes x, y, z. They can be written together as

$$\widehat{s}_i\psi_k = -\mathrm{i}e_{ikl}\psi_l.$$

The complex vector $\boldsymbol{\psi}$ can be put in the form $\boldsymbol{\psi} = \mathrm{e}^{\mathrm{i}\alpha}(\boldsymbol{u} + \mathrm{i}\boldsymbol{v})$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are real vectors, which can be taken to be mutually perpendicular if the common phase α is suitably chosen. The two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ determine a plane which has the property that the spin component perpendicular to it can take only the values ± 1.

§ 58. The operator of finite rotations

Let us now return to the transformation of spinors, and show how the coefficients of this transformation can in fact be expressed in terms of the angles of rotation of the coordinate axes.

By the definition of the angular momentum operator (in this case, the spin operator), $1 + \mathrm{i}\delta\varphi \cdot \boldsymbol{n}\widehat{\boldsymbol{s}}$ is the operator of a rotation through an angle $\delta\varphi$ about a direction specified by the unit vector $\boldsymbol{n}$; for application to the wave function of a particle with spin 1/2, i.e. a spinor of rank one, we must take $\widehat{\boldsymbol{s}} = \widehat{\boldsymbol{\sigma}}/2$ in this operator. The operator of a rotation through a finite angle φ about the same direction will be correspondingly given by

$$\widehat{U}_{\boldsymbol{n}} = \exp\left(\mathrm{i}\varphi\boldsymbol{n}\widehat{\boldsymbol{\sigma}}/2\right) \tag{58.1}$$

cf. (15.13). Like any function of the Pauli matrices (see §55, Problem 1), this expression reduces to one that is linear in these matrices:

$$\widehat{U}_{\boldsymbol{n}} = \cos(\varphi/2) + \mathrm{i}\boldsymbol{n}\widehat{\boldsymbol{\sigma}} \cdot \sin(\varphi/2). \tag{58.2}$$

For example, with a rotation about the z-axis,

$$\widehat{U}_z(\varphi) = \cos\frac{\varphi}{2} + i\widehat{\sigma}_z \sin\frac{\varphi}{2} = \begin{pmatrix} \exp(i\varphi/2) & 0 \\ 0 & \exp(-i\varphi/2) \end{pmatrix} \qquad (58.3)$$

This means that the components of the spinor are transformed in such a rotation according to

$$\psi^{1'} = \psi^1 e^{i\varphi/2}, \quad \psi^{2'} = \psi^2 e^{-i\varphi/2}.$$

In particular, in a rotation through an angle 2π the spinor components change sign; spinors of any odd rank must therefore have the same property (cf. the end of §55).

Similarly, we can find the matrices of transformations consisting of a rotation through an angle φ about the x-axis or the y-axis:

$$\widehat{U}_x(\varphi) = \begin{pmatrix} \cos(\varphi/2) & i\sin(\varphi/2) \\ i\sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix}, \widehat{U}_y(\varphi) = \begin{pmatrix} \cos(\varphi/2) & \sin(\varphi/2) \\ -\sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix}.$$
$$(58.4)$$

We may note the particular case of a rotation through an angle π about the y-axis, for which

$$\psi^{1'} = \psi^2, \quad \psi^{2'} = -\psi^1,$$

i.e.

$$\psi^{1'} = \psi_1, \quad \psi^{2'} = \psi_2. \qquad (58.5)$$

It is now easy to write down the transformation matrix for any rotation of the coordinate axes, as a function of the Eulerian angles which specify the rotation.

A rotation of the axes, defined by the Eulerian angles α, β, γ, is carried out in three stages: (1) a rotation through the angle α $(0 \leqslant \alpha \leqslant 2\pi)$ about the z-axis, (2) a rotation through the angle β $(0 \leqslant \beta \leqslant \pi)$ about the new position of the y-axis (ON in Fig. 9, called the *line of nodes*), (3) a rotation through the angle γ $(0 \leqslant \gamma \leqslant 2\pi)$ about the resulting final position z' of the z-axis.[19])

---

[19]) The systems xyz and $x'y'z'$ are, as always, right-handed, and a positive angle corresponds to the movement of a corkscrew advanced in the positive direction of the axis of rotation.

The definition of the Eulerian angles given here (and usual in quantum mechanics) differs from that in *Mechanics*, §35, in that the second rotation is about the y-axis and not the x-axis. The angles α, β, γ are related to the angles φ, θ, ψ used in *Mechanics* (*not* the spherical polar angles φ and θ) by

$$\varphi = \alpha + \frac{\pi}{2}, \quad \theta = \beta, \quad \psi = \gamma - \frac{\pi}{2}.$$
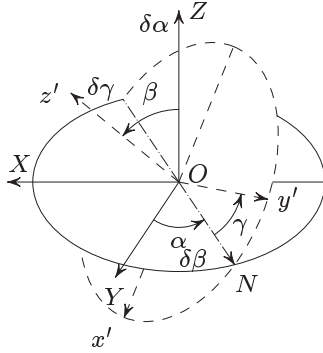
FIG.20

It is evident that the angles α and β are the spherical polar angles φ and θ of the new z'-axis with respect to the xyz axes: $\alpha = \varphi, \beta = \theta$.

In accordance with this manner of rotating the axes, the matrix of the complete transformation is equal to the product of three matrices (58.3)and (58.4):

$$\widehat{U}(\alpha,\beta,\gamma) = \widehat{U}_z(\gamma)\widehat{U}_y(\beta)\widehat{U}_z(\alpha)$$

By direct multiplication of the matrices we finally obtain

$$\widehat{U}(\alpha,\beta,\gamma) = \begin{pmatrix} \cos(\beta/2)\mathrm{e}^{\mathrm{i}(\alpha+\gamma)/2} & \sin(\beta/2)\mathrm{e}^{-\mathrm{i}(\alpha-\gamma)/2} \\ -\sin(\beta/2)\mathrm{e}^{\mathrm{i}(\alpha-\gamma)/2} & \cos(\beta/2)\mathrm{e}^{-\mathrm{i}(\alpha+\gamma)/2} \end{pmatrix}. \tag{58.6}$$

Spinors of higher ranks are, by definition, transformed as products of components of a spinor of rank one. In physical applications, however, we are interested in the wave functions ψ_{jm} rather than the transformation laws of the spinors themselves.

Let the functions ψ_{jm} $(m = j, j-1, \ldots, -j)$ describe, in a coordinate system xyz, a state having a definite value of the angular momentum j, and ψ_{jm} the same state for the axes $x'y'z'$; in the first case m is the value of j_z, and in the second case m' is the value of $j_{z'}$. The two sets of functions are connected by linear relations, which we write in the form

$$\psi_{jm} = \sum_{m'} D_{m'm}^{(f)}(\alpha,\beta,\gamma)\,\psi_{jm'} \tag{58.7}$$

The coefficients $D_{m'm}^{(j)}$ form a matrix of order $2j+1$ with respect to m' and m, called the *finite-rotation matrix* $\widehat{D}^{(j)}$; its elements are functions of the angles α, β, γ of rotation of the system $x'y'z'$ relative to xyz.

The finite-rotation matrix can be built up by means of the spinor representation of the functions ψ_{jm}. For $j = 1/2$, the two functions $\psi_{1/2m}(m = \pm 1/2)$ form a covariant spinor of rank 1. According to (56.13), its transformation from $x'y'z'$ to xyz is effected by the matrix $\widehat{U}$ (58.6), so that $\widehat{D}^{(1/2)} = \widehat{U}$.[20] Its elements may be written

$$D_{m'm}^{1/2} = \mathrm{e}^{\mathrm{i}m'\gamma}d_{m'm}(\beta)\mathrm{e}^{\mathrm{i}m\alpha}$$

---

[20] Note that the matrix indices in (58.7) are placed in the order that corresponds to multiplying the columns of the matrix $\widehat{D}^{(j)}$ by the functions $\psi_{jm'}$ arranged in a row. In the symbolic notation, (58.7) would have to be written $\psi_{jm} = (\psi'_j\widehat{D}^{(j)})_m$ in accordance with (56.13).

where

$$d_{m'm}^{1/2}(\beta) = \begin{array}{c|cc} \diagdown^{\textstyle m} & 1/2 & -1/2 \\ \hline m' & & \\ \hline 1/2 & \cos(\beta/2) & \sin(\beta/2), \\ -1/2 & -\sin(\beta/2) & \cos(\beta/2). \end{array} \qquad (58.8)$$

For any value of j, the functions ψ_{jm} are related to the components of a symmetrical covariant spinor of rank $2j$ by (57.6). The transformation matrix for the components of a spinor of rank $2j$ is the product of $2j$ matrices $\widehat{D}^{(1/2)}$, each acting on one of the spinor indices. Carrying out the multiplication and returning to the functions ψ_{jm}, we find their transformation matrix:

$$D_{m'm}^{(j)}(\alpha, \beta, \gamma) = e^{im'\gamma} d_{m'm}^{(j)}\beta e^{im\alpha}, \qquad (58.9)$$

the functions $d_{m'm}^{(j)}(\beta)$ being given by[21]

$$d_{m'm}^{(j)}(\beta) = \left[\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!}\right]^{1/2} \left(\cos\frac{\beta}{2}\right)^{m'+m} \times$$
$$\times \left(\sin\frac{\beta}{2}\right)^{m'-m} P_{j-m'}^{(m'-m,m'+m)}(\cos\beta), \quad (58.10)$$

where

$$P_n^{(a,b)}(\cos\beta) = \frac{(-1)^n}{2^n n!}(1-\cos\beta)^{-\alpha}(1+\cos\beta)^{-b} \times$$
$$\times \left(\frac{d}{d\cos\beta}\right)^n \left[(1-\cos\beta)^{a+n}(1+\cos\beta)^{b+n}\right] \quad (58.11)$$

are called *Jacobi polynomials*.[22] We may note that

$$P_n^{(a,b)}(-\cos\beta) = (-1)^n P_n^{(b,a)}(\cos\beta). \qquad (58.12)$$

The functions $d_{m'm}^{(j)}$ possess a number of symmetry properties which might be derived from the expressions (58.11) and (58.12), but it is simpler to obtain them directly from the definition as coefficients in the rotational transformation.

---

[21] The calculations are described by A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton, 1957. The definition of the functions $d_{m'm}^{(j)}$ by (58.9) differs from that used in Edmonds's book by the interchange of α and γ, this being the more natural treatment in the approach given here.

[22] See §e of the Mathematical Appendices, formula (e.11), for the relation between these polynomials and the hypergeometric series.

The matrix $\widehat{D}^{(j)}$ is unitary, being the matrix of a rotational transformation. Since the transformation inverse to the rotation (α, β, γ) is the rotation $(-\gamma, -\beta, -\alpha)$, we have for the real matrix $\widehat{d}^{(j)}$ the relations

$$d^{(j)}_{m'm}(-\beta) = d^{(j)}_{m'm}(\beta) \tag{58.13}$$

The following equations are also valid:

$$d^{(j)}_{m'm}(\beta) = d^{(j)}_{-m,-m'}(\beta) \tag{58.14}$$

$$d^{(j)}_{m'm}(\pi) = (-1)^{j+m}\delta_{m',-m},$$
$$d^{(j)}_{m'm}(-\pi) = (-1)^{j-m}\delta_{m',-m}, \quad d^{(j)}_{m'm}(0) = \delta_{m'm}. \tag{58.15}$$

When $j = 1/2$ these are evident from (58.8); the generalization to arbitrary j is evident from the manner of construction of the transformation matrix, described above.

A rotation through an angle $\pi - \beta$ can be carried out as two successive rotations through ψ and $-\beta$:

$$d^{(j)}_{m'm}(\pi - \beta) = \sum_{m''} d^{(j)}_{m'm''}(\pi) d^{(j)}_{m''m}(-\beta) = (-1)^{j-m'} d^{(j)}_{-m'm}(-\beta),$$

or, using (58.13),

$$d^{(j)}_{m'm}(\pi - \beta) = (-1)^{j-m'} d^{(j)}_{m,-m'}(-\beta). \tag{58.16}$$

The result of two rotations about the same axis is independent of the sequence in which they occur. We must therefore arrive at the same result by carrying out the rotations through $-\beta$ and π in the opposite order. Comparison of the result with (58.16) gives the relation

$$d^{(j)}_{m'm}(\beta) = (-1)^{m'-m} d^{(j)}_{-m',-m}(\beta). \tag{58.17}$$

From (58.17), (58.14) and (58.13), it follows that

$$d^{(j)}_{m'm}(\beta) = (-1)^{m'-m} d^{(j)}_{mm'}(\beta) = (-1)^{m'-m} d^{(j)}_{m'm}(-\beta). \tag{58.18}$$

Using (58.13)–(58.18), we can deduce various symmetry properties of the complete matrix elements $D^{(j)}_{m'm}$. In particular, the complex conjugate function is given by

$$D^{(j)*}_{m'm}(\alpha, \beta, \gamma) = D^{(j)}_{m'm}(-\alpha, \beta, -\gamma) = (-1)^{m'-m} D^{(j)}_{-m',-m}(\alpha, \beta, \gamma). \tag{58.19}$$

Mathematically, the matrices $\widehat{D}^{(j)}$ give the unitary irreducible representations of the rotation group having dimension $2j+1$ (see §98 below). Hence we have immediately the orthonormality relation

$$\int D_{m'_1 m_1}^{(j_1)*}(\alpha, \beta, \gamma) D_{m'_2 m_2}^{(j_2)}(\alpha, \beta, \gamma) \frac{d\omega}{8\pi^2} = \frac{1}{2j_1 + 1} \delta_{j_1 j_2} \delta_{m_1 m_2} \delta_{m'_1 m'_2}, \quad (58.20)$$

where $d\omega = \sin\beta d\alpha d\beta d\gamma$.

The orthogonality of the functions with respect to the suffixes m and m' is ensured by the factor $\exp\{i(m\alpha + m'\gamma)\}$; that with respect to the index j arises from the functions $d_{m'm}^{(j)}$, for which we have

$$\int_0^\pi d_{m'm}^{(j_1)}(\beta) d_{m'm}^{(j_2)}(\beta) \frac{\sin\beta d\beta}{2} = \frac{1}{2j_1 + 1} \delta_{j_1 j_2}. \quad (58.21)$$

Lastly, we shall give for reference the expressions for the functions for various particular values of the parameters. For $j = 1$, we have

$$
d_{m'm}^{(1)}(\beta) =
\begin{array}{c|ccc}
\diagbox{m'}{m} & 1 & 0 & -1 \\
\hline
1 & \frac{1}{2}(1 + \cos\beta) & \frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1 - \cos\beta), \\
0 & -\frac{1}{\sqrt{2}}\sin\beta & \cos\beta & \frac{1}{\sqrt{2}}\sin\beta, \\
-1 & \frac{1}{2}(1 - \cos\beta) & -\frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1 + \cos\beta).
\end{array}
\quad (58.22)
$$

For integral $j = l$ and $m' = 0$, formulae (58.10) and (58.11) give

$$d_{0m}^{(l)}(\beta) = (-1)^m d_{m0}^{(l)}(\beta) = (-1)^m \sqrt{\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\beta). \quad (58.23)$$

The derivation of this formula is easily seen from the original definition (58.7). We shall assign the values of the functions $\psi_{jm'}$. on the right of (58.7) to the z-axis, on which (for $j = l$)

$$Y_{lm'}(\boldsymbol{n}_{z'}) = i^l \sqrt{\frac{2l+1}{4\pi}} \delta_{m'0}. \quad (58.24)$$

The function ψ_{jm} on the left is then the spherical harmonic function $Y_{lm}(\beta, \alpha)$ of the spherical polar angles $\varphi \equiv \alpha$, $\theta \equiv \beta$ giving the direction of the z'-axis. Substitution of (58.24) in (58.7) leads to

$$Y_{lm}(\beta, \alpha) = i^l \sqrt{\frac{2l+1}{4\pi}} D_{0m}^{(l)}(\alpha, \beta, \gamma), \quad (58.25)$$

which is equivalent to (58.23).

Lastly, there is the following expression for the function with the maximum possible value of m or m':

$$d_{jm}^{(j)}(\beta) = (-1)^{j-m} d_{mj}^{(j)} =$$

$$= \left[\frac{(2j)!}{(j+m)!(j-m)!}\right]^{1/2} \left(\cos\frac{\beta}{2}\right)^{j+m} \left(\sin\frac{\beta}{2}\right)^{j-m}. \quad (58.26)$$

§ 59. Partial polarization of particles

By a suitable choice of the direction of the z-axis, we can always cause one component (e.g. ψ^2) of a given spinor ψ^λ, the wave function of a particle with spin $1/2$, to vanish. This is evident from the fact that a direction in space is determined by two quantities (angles), i.e. the number of disposable parameters is just equal to the number of quantities (the real and imaginary parts of the complex ψ^2) which it is desired to make zero.

Physically this means that, if a particle with spin $1/2$ (for definiteness, we shall speak of an electron) is in a state described by a spin wave function, then there is a direction in space in which the component of the particle spin has the definite value $\sigma = 1/2$. We can say that in such a state the electron is *completely polarized.*

There are also, however, states of an electron which may be said to be *partially polarized.* Such states are not described by wave functions but only by density matrices, i.e. they are mixed states (with respect to spin) (see §14).

The spin or *polarization* density matrix of an electron is a spinor $\rho^{\lambda\mu}$ of rank two normalized by the condition

$$\rho^\lambda{}_\lambda = \rho^1{}_1 + \rho_2^2 = 1 \quad (59.1)$$

and satisfying the "Hermitian" condition

$$(\rho^\lambda{}_\mu)^* = \rho^\mu{}_\lambda. \quad (59.2)$$

For a pure (i.e. completely polarized) spin state of the electron the spinor $\rho^\lambda{}_\mu$ reduces to a product of components of the wave function ψ^λ:

$$\rho^\lambda{}_\mu = \psi^\lambda(\psi^\mu)^*. \quad (59.3)$$

The diagonal components of the density matrix determine the probabilities of the values $+1/2$ and $-1/2$ of the z-component of the electron spin. The mean value of this component is therefore

$$\overline{s}_z = \frac{1}{2}(\rho^1{}_1 - \rho^2{}_2),$$

or, using (59.1),

$$\rho^1{}_1 = \frac{1}{2} + \overline{s}_z, \quad \rho^2{}_2 = \frac{1}{2} - \overline{s}_z. \tag{59.4}$$

In a pure state the mean value of the quantities $s_\pm = s_x \pm is_y$ is calculated as

$$\overline{s}_+ = \psi^{\lambda*}\widehat{s}_+\psi^\lambda, \quad \overline{s}_- = \psi^{\lambda*}\widehat{s}_-\psi^\lambda.$$

Since, according to (55.6) and (55.7), the operators $\widehat{s}_\pm$ are given by the matrices

$$\overline{s}_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \overline{s}_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

we find that

$$\overline{s}_+ = \psi^{1*}\psi^2, \quad \overline{s}_-\psi^{2*}\psi^1.$$

Accordingly we have in a mixed state

$$\rho^1{}_2 = \overline{s}_-, \quad \rho^2{}_1 = \overline{s}_+. \tag{59.5}$$

Using the Pauli matrices, formulae (59.4) and (59.5) can be combined as

$$\rho^\lambda{}_\mu = \frac{1}{2}(\delta^\lambda{}_\mu + 2\widehat{\boldsymbol{\sigma}}^\lambda{}_\mu\overline{\boldsymbol{s}}). \tag{59.6}$$

Thus all the components of the polarization density matrix of the electron are expressed in terms of the mean values of components of its spin vector. In other words, the real vector $\overline{\boldsymbol{s}}$ entirely determines the polarization properties of a particle with spin 1/2. In the limit of complete polarization one of the components of this vector (with an appropriate choice of the directions of the axes) is 1/2 and the other two are zero. In the opposite case of an unpolarized state all three components are zero. In the general case of an arbitrary partial polarization and any choice of the coordinate system we have $0 \leqslant \rho \leqslant 1$, where

$$\rho = 2(\overline{s}_x^2 + \overline{s}_y^2 + \overline{s}_z^2)^{1/2}$$

is a quantity which may be called the degree of polarization of the electron.

For a particle of arbitrary spin s, the density matrix is a spinor $\rho^{\lambda\mu\cdots}{}_{\rho\sigma\cdots}$ of rank $4s$, symmetrical in the first $2s$ and the last $2s$ indices and satisfying the conditions

$$\rho^{\lambda\mu\cdots}{}_{\lambda\mu\cdots} = 1, \tag{59.7}$$

$$\left(\rho^{\lambda\mu\cdots}{}_{\rho\sigma\cdots}\right)^* = \rho^{\rho\sigma\cdots}{}_{\lambda\mu\cdots}. \tag{59.8}$$

To calculate the number of independent components of the density matrix, we note that, among the possible sets of values of the indices $\lambda, \mu, \ldots$ (or $\rho, \sigma, \ldots$) there are only $2s + 1$ which are essentially different. Using also

the fact that the components of the spinor $\rho^{\lambda\mu\cdots}{}_{\rho\sigma\cdots}$ are related by (59.7), we find that the number of different components is $(2s+1)^2 - 1 = 4s(s+1)$. Although these components are complex, the relation (59.8) shows that this does not increase the total number of independent quantities describing the state of partial polarization of the particle, which is therefore $4s(s+1)$.[23] For comparison, it may be remarked that the state of complete polarization of the particle is described by only $4s$ quantities (the $2s+1$ complex components of the wave function $\psi^{\lambda\mu\cdots}$, related by one normalization condition and containing one common phase which is unimportant in the description of the state).

Like any spinor of rank $4s$, the spinor $\rho^{\lambda\mu\cdots}{}_{\rho\sigma\cdots}$ is equivalent to a set of irreducible tensors of ranks $4s, 4s - 2, \ldots, 0$. In the present case there is only one tensor of each rank, since, on account of the symmetry properties of the spinor $\rho^{\lambda\mu\cdots}{}_{\rho\sigma\cdots}$, each contraction of it can be carried out in only one way: with respect to any one of the indices $\lambda, \mu, \ldots$, and one of $\rho, \sigma, \ldots$. In addition, the scalar (tensor of rank 0) does not appear, reducing to unity by virtue of the condition (59.7).

§ 60. Time reversal and Kramers' theorem

The symmetry of motion with respect to a change in the sign of the time is expressed in quantum mechanics by the fact that, if ψ is the wave function of a stationary state of the system, the "time-reversed" wave function (which we denote by ψ^{rev}) describes a possible state with the same energy. At the end of §18 it has been pointed out that ψ^{rev} is the same as the complex conjugate function ψ^*. In this simple form the statement applies to wave functions where the spin of particles is neglected. When spin is present, a refinement is necessary.

Let us take the wave function of a particle of spin s in the form of the contravariant spinor $\psi^{\lambda\mu\cdots}$ (of rank $2s$). On taking the complex conjugate function $\psi^{\lambda\mu\cdots*}$ we obtain a set of quantities which are transformed as components of a covariant spinor. Hence the operation of time reversal corresponds to a change from the wave function $\psi^{\lambda\mu\cdots}$ to a new wave function whose covariant components are given by

$$\psi^{\mathrm{rev}}_{\lambda\mu\cdots} = \psi^{\lambda\mu\cdots*}. \tag{60.1}$$

For a given set of values of the indices $\lambda, \mu, \ldots$, the components of covari-

---

[23] When these quantities are given, so are the mean values of the components of the vector $\boldsymbol{s}$ and all their powers and products $2, 3, \ldots, 2s$ at a time, which do not reduce to lower powers (see §55, Problem 3).

ant and contravariant spinors correspond to values of the angular-momentum component which differ in sign. In terms of the functions $\psi_{s\sim}$ therefore, time reversal corresponds to a change from $\psi_{s\sigma}$ to $\psi_{s,-\sigma}$, as it should, since a change in the sign of the time changes the direction of the angular momentum. The exact relation is given by (60.1):

$$\psi_{s,-\sigma}^{\text{rev}} = \psi_{s\sigma}^{*}(-1)^{s-\sigma}. \tag{60.2}$$

Thus the change $\psi_{s\sigma} \to \psi_{s\sigma}^{*}$ required by the operation of time reversal signifies the change[24]

$$\psi_{s\sigma} \to \psi_{s,-\sigma}(-1)^{s-\sigma} \tag{60.3}$$

When this operation is repeated, we have

$$\psi_{s\sigma} \to \psi_{s,-\sigma}(-1)^{s-\sigma} \to \psi_{s\sigma}(-1)^{s-\sigma}(-1)^{s+\sigma} = \psi_{s\sigma}(-1)^{2s}.$$

Thus a twofold time reversal restores the wave function to its original value only if the spin is integral; if the spin is half-integral, the sign of the wave function is changed.

Let us consider an arbitrary system of interacting particles. The orbital and spin angular momenta of such a system are not in general separately conserved when relativistic interactions are taken into account. Only the total angular momentum $\boldsymbol{J}$ is conserved. If there is no external field, each energy level of the system has $(2J + 1)$-fold degeneracy. When an external field is applied, the degeneracy is removed. The question arises whether the degeneracy can be removed completely, i.e. so that the system has only simple levels. This is closely related to the symmetry with respect to time reversal.

In classical electrodynamics the equations are invariant with respect to a change in the sign of the time, if the electric field is left unchanged and the sign of the magnetic field is reversed.[25] This fundamental property of motion must be preserved in quantum mechanics. Hence, not only in a closed system but in any external electric field (there being no magnetic field), there is symmetry with respect to time reversal.

The wave functions of the system are spinors $\psi^{\lambda\mu\cdots}$, whose rank n is twice the sum of the spins sa of all the particles ($n = 2\sum s_a$); this sum may not be equal to the total spin S of the system.

According to what was said above, we can assert that, in any electric field, the wave function and its time reversal must correspond to states with

---

[24] Note that the rule for the complex conjugate of a spherical harmonic function, according to (28.9), coincides with the general rule (60.3).

[25] See, for example, *Fields*, §17, and the end of §111 below.

the same energy. If a level is non-degenerate, it is necessary that these states should be identical, i.e. the corresponding wave functions must be the same apart from a constant factor (both, of course, being expressed as similar (covariant or contravariant) spinors).

We write $\psi^{\text{rev}}_{\lambda\mu\ldots} = C\psi_{\lambda\mu\ldots}$ or, by (60.1),

$$\psi^{\lambda\mu\cdots*} = C\psi_{\lambda\mu\ldots}, \tag{60.4}$$

where C is a constant.

Taking the complex conjugate of both sides of this equation, we obtain

$$\psi^{\lambda\mu\ldots} = C^*\psi^*_{\lambda\mu\ldots}.$$

We lower the indices on the left-hand side of the equation and correspondingly raise them on the right. This means that we multiply both sides of the equation by $g_{\alpha\lambda}g_{\beta\mu\ldots}$ and sum over the indices $\lambda, \mu, \ldots$; on the right-hand side we must use the fact that

$$g_{\alpha\lambda}g_{\beta\mu\ldots} = (-1)^n g^{\lambda\alpha}g^{\mu\beta}\ldots$$

As a result we have

$$\psi_{\lambda\mu\ldots} = C^*(-1)^n\psi^{\lambda\mu\cdots*}.$$

Substituting $\psi^{\lambda\mu\cdots*}$ from (60.4), we find

$$\psi_{\lambda\mu\ldots} = (-1)^n CC^*\psi_{\lambda\mu\ldots}.$$

This equation must be satisfied identically, i.e. we must have $(-1)^n CC^* = 1$. Since, however, $|C|^2$ is always positive, it is clear that this is possible only for even n (i.e. for integral values of the sum $\sum s_a$). For odd n (half-integral values of $\sum s_a$) the condition (60.4) cannot be fulfilled.[26]

Thus we reach the result that an electric field can completely remove the degeneracy only for a system with an integral value of the sum of the spins of the particles. For a system with a half-integral value of this sum, in an arbitrary electric field, all the levels must be doubly degenerate, and complex conjugate spinors correspond to two different states with the same energy[27] (H. A. Kramers 1930).

One further, mathematical, comment may be made. A relation of the form (60.4) with a real constant C is mathematically the condition that the

---

[26] When the sum $\sum s_a$ is integral (or halt-integral), all possible values of the total spin S of the system are also integral (or half-integral).

[27] If the electric field possesses a high (cubic) symmetry, fourfold degeneracy may occur (see §99, including the Problem).

components of the spinor may be put in correspondence with a set of real quantities, and may be called the condition for the spinor to be "real".[28] The impossibility of fulfilling the condition (60.4) for odd n signifies that no real quantity can correspond to a spinor of odd rank. For even n, on the other hand, the condition (60.4) can be satisfied, and C can be real. In particular, a real vector can correspond to a symmetrical spinor of rank two if the condition (60.4) is satisfied with $C = 1$:

$$\psi^{\lambda\mu*} = \psi_{\lambda\mu}$$

(as is easily seen by means of (57.8) and (57.9)). The condition (60.4) with $C = 1$ is in fact the condition for a symmetrical spinor of any even rank to be "real".

---

[28] It is meaningless to call the spinor real in the literal sense, since complex conjugate spinors have different laws of transformation.

CHAPTER IX

IDENTITY OF PARTICLES

§ 61. The principle of indistinguishability of similar particles

IN classical mechanics, identical particles (electrons, say) do not lose their "individuality", despite the identity of their physical properties. For we can imagine the particles at some instant to be "numbered", and follow the subsequent motion of each of these in its path; then at any instant the particles can be identified.

In quantum mechanics the situation is entirely different. We have already mentioned several times that, by virtue of the uncertainty principle, the concept of the path of an electron ceases to have any meaning. If the position of an electron is exactly known at a given instant, its coordinates have no definite values even at the next instant. Hence, by localizing and numbering the electrons at some instant, we make no progress towards identifying them at subsequent instants; if we localize one of the electrons, at some other instant, at some point in space, we cannot say which of the electrons has arrived at this point.

Thus, in quantum mechanics, there is in principle no possibility of separately following each of a number of similar particles and thereby distinguishing them. We may say that, in quantum mechanics, identical particles entirely lose their "individuality". The identity of the particles with respect to their physical properties is here very far-reaching: it results in the complete indistinguishability of the particles.

This principle of the *indistinguishability of similar particles*, as it is called, plays a fundamental part in the quantum theory of systems composed of identical particles. Let us start by considering a system of only two particles. Because of the identity of the particles, the states of the system obtained from each other by merely interchanging the two particles must be completely equivalent physically. This means that, as a result of this interchange, the wave function of the system can change only by an unimportant phase factor.

Let $\psi(\xi_1, \xi_2)$ be the wave function of the system, ξ_1 and ξ_2 conventionally denoting the three coordinates and the spin projection for each particle. Then we must have

$$\psi(\xi_1, \xi_2) = \mathrm{e}^{\mathrm{i}\alpha}\psi(\xi_2, \xi_1),$$

where α is some real constant. By repeating the interchange, we return to the original state, while the function ψ is multiplied by $\mathrm{e}^{2\mathrm{i}\alpha}$. Hence it follows that $\mathrm{e}^{2\mathrm{i}\alpha} = 1$, or $\mathrm{e}^{\mathrm{i}\alpha} = \pm 1$. Thus

$$\psi(\xi_1, \xi_2) = \pm\psi(\xi_2, \xi_1).$$

We thus reach the result that there are only two possibilities: the wave function is either *symmetrical* (i.e. it is unchanged when the particles are interchanged) or *antisymmetrical* (i.e. it changes sign when this interchange is made). It is obvious that the wave functions of all the states of a given system must have the same symmetry; otherwise, the wave function of a state which was a superposition of states of different symmetry would be neither symmetrical nor antisymmetrical.

This result can be immediately generalized to systems consisting of any number of identical particles. For it is clear from the identity of the particles that, if any pair of them has the property of being described by, say, symmetrical wave functions, any other pair of such particles has the same property. Hence the wave function of identical particles must either be unchanged when any pair of particles are interchanged (and hence when the particles are permuted in any manner), or change sign when any pair are interchanged. In the first case we speak of a *symmetrical* wave function, and in the second case of an *antisymmetrical* one.

The property of being described by symmetrical or antisymmetrical wave functions depends on the nature of the particles. Particles described by antisymmetrical functions are said to obey *Fermi–Dirac statistics* (or to be *fermions*), while those which are described by symmetrical functions are said to obey *Bose-Einstein statistics* (or to be *bosons*).[1]

From the laws of relativistic quantum mechanics it can be shown (see RQT, §25) that the statistics obeyed by particles is uniquely related to their spin: particles with half-integral spin are fermions, and those with integral spin are bosons.

---

[1] This terminology refers to the statistics which describes a perfect gas composed of particles with antisymmetrical and symmetrical wave functions respectively. In actual fact we are concerned here not only with a different statistics, but essentially with a different mechanics. Fermi statistics was proposed by E. Fermi for electrons in 1926, and its relation to quantum mechanics was elucidated by P. A. M. Dirac (1926). Bose statistics was proposed by S. N. Bose for light quanta, and generalized by A. Einstein (1924).

The statistics of complex particles is determined by the parity of the number of elementary fermions entering into their composition. For an interchange of two identical complex particles is equivalent to the simultaneous interchange of several pairs of identical elementary particles. The interchange of bosons does not change the wave function, while the interchange of fermions changes its sign. Hence complex particles containing an odd number of elementary fermions obey Fermi statistics, while those containing an even number obey Bose statistics. This result is, of course, in agreement with the above rule, since a complex particle has an integral or a half-integral spin according as the number of particles with half-integral spin entering into its composition is even or odd.

Thus atomic nuclei of odd atomic weight (i.e. containing an odd number of neutrons and protons) obey Fermi statistics, and those of even atomic weight obey Bose statistics. For atoms, which contain both nuclei and electrons, the statistics is evidently determined by the parity of the sum of the atomic weight and the atomic number.

Let us consider a system composed of N identical particles, whose mutual interaction can be neglected. Let $\psi_1, \psi_2, \ldots$ be the wave functions of the various stationary states which each of the particles separately may occupy. The state of the system as a whole can be defined by giving the numbers of the states which the individual particles occupy. The question arises how the wave function ψ of the whole system should be constructed from the functions $\psi_1, \psi_2, \ldots$.

Let $p_1, p_2, \ldots, p_N$ be the numbers of the states occupied by the individual particles (some of these numbers may be the same). For a system of bosons, the wave function $\psi(\xi_1, \xi_2, \ldots, \xi_N)$ is given by a sum of products of the form

$$\psi_{p_1}(\xi_1)\psi_{p_2}(\xi_2)\ldots\psi_{p_n}(\xi_N), \tag{61.1}$$

with all possible permutations of the different suffixes $p_1, p_2, \ldots$; this sum clearly possesses the required symmetry property. For example, for a system of two particles in different states ($p_1 \neq p_2$),

$$\psi(\xi_1, \xi_2) = \frac{1}{\sqrt{2}}\left[\psi_{p_1}(\xi_1)\psi_{p_2}(\xi_2) + \psi_{p_1}(\xi_2)\psi_{p_2}(\xi_1)\right]. \tag{61.2}$$

The factor $1/\sqrt{2}$ is introduced for normalization purposes; all the functions $\psi_1, \psi_2, \ldots$ are orthogonal and are supposed normalized.

In the general case of a system containing an arbitrary number N of particles, the normalized wave function is

$$\psi_{N_1 N_2 \ldots} = \left(\frac{N_1! N_2! \ldots}{N!}\right)^{1/2} \sum \psi_{p_1}(\xi_1)\psi_{p_2}(\xi_2)\ldots\psi_{p_N}(\xi_N), \tag{61.3}$$

where the sum is taken over all permutations of the different suffixes $p_1, p_2, \ldots, p_N$ and the numbers N_i show how many of these suffixes have the same value i (with $\sum N_i = N$). In the integration of $|\psi|^2$ over $\xi_1, \xi_2, \ldots, \xi_N$, all terms vanish except the squared modulus of each term of the sum;[2] since the total number of terms in the sum (61.3) is evidently $N!/(N_1!N_2!\ldots)$, the normalization factor in (61.3) is obtained.

For a system of fermions, the wave function ψ is an antisymmetrical combination of the products (61.1). For a system of two particles we have

$$\psi(\xi_1, \xi_2) = \frac{1}{\sqrt{2}} \left[\psi_{p_1}(\xi_1)\psi_{p_2}(\xi_2) - \psi_{p_1}(\xi_2)\psi_{p_2}(\xi_1) \right]. \tag{61.4}$$

For the general case of N particles, the wave function can be written in the form of a determinant

$$\psi_{N_1 N_2 \ldots} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{p_1}(\xi) & \psi_{p_1}(\xi_2) & \ldots & \psi_{p_1}(\xi_N) \\ \psi_{p_2}(\xi) & \psi_{p_2}(\xi_2) & \ldots & \psi_{p_2}(\xi_N) \\ \ldots & \ldots & \ldots & \ldots \\ \psi_{p_N}(\xi) & \psi_{p_N}(\xi_2) & \ldots & \psi_{p_N}(\xi_N) \end{vmatrix}. \tag{61.5}$$

Here an interchange of two particles corresponds to an interchange of two columns of the determinant, as a result of which the latter changes sign.

The following important result is a consequence of the expression (61.5). If among the numbers $p_1, p_2, \ldots$ two are the same, two rows of the determinant are the same, and it therefore vanishes identically. It will be different from zero only when all the numbers $p_1, p_2, \ldots$ are different. Thus, in a system consisting of identical fermions, no two (or more) particles can be in the same state at the same time. This is called *Pauli's principle* (1925).

§ 62. Exchange interaction

The fact that Schrödinger's equation does not take account of the spin of particles does not invalidate this equation or the results obtained by means of it. This is because the electrical interaction of the particles does not depend on their spins.[3] Mathematically, this means that the Hamiltonian of a system of electrically interacting particles (in the absence of a magnetic field) does not contain the spin operators, and hence, when it is applied to

---

[2] The integration over ξ is conventionally understood in §§63–65 as including integration over the coordinates and summation over σ.

[3] This is true only so long as we consider the non-relativistic approximation. When relativistic effects are taken into account, the interaction of charged particles does depend on their spin.

the wave function, it has no effect on the spin variables. Hence Schrödinger's equation is actually satisfied by each component of the wave function; in other words, the wave function of the system of particles can be written in the form of a product

$$\psi(\xi_1, \xi_2) = \chi(\sigma_1, \sigma_2, \dots)\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots),$$

where the function φ depends only on the coordinates of the particles and the function χ only on their spins. We call the former a *coordinate* or *orbital* wave function, and the latter a *spin* wave function. Schrödinger's equation essentially determines only the coordinate function φ, the function χ remaining arbitrary. In any instance where we are not interested in the actual spin of the particles, we can therefore use Schrödinger's equation and regard as the wave function the coordinate function alone, as we have done hitherto.

However, despite the fact that the electrical interaction of the particles is independent of their spin, there is a peculiar dependence of the energy of the system on its total spin, arising ultimately from the principle of indistinguishability of similar particles.

Let us consider a system consisting of only two identical particles. By solving Schrödinger's equation we find a series of energy levels, to each of which there corresponds a definite symmetrical or antisymmetrical coordinate wave function $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$. For, by virtue of the identity of the particles, the Hamiltonian (and therefore the Schrödinger's equation) of the system is invariant with respect to interchange of the particles. If the energy levels are not degenerate, the function $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ can change only by a constant factor when the coordinates $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are interchanged; repeating this interchange, we see that this factor can only be[4] ± 1.

Let us first suppose that the particles have zero spin. The spin factor for such particles is absent altogether, and the wave function reduces to the coordinate function $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$, which must be symmetrical (since particles with zero spin obey Bose statistics). Thus not all the energy levels obtained by a formal solution of Schrödinger's equation can actually exist; those to which antisymmetrical functions φ correspond are not possible for the system under consideration.

The interchange of two similar particles is equivalent to the operation of inversion of the coordinate system (the origin being taken to bisect the line joining the two particles). On the other hand, the result of inversion is to multiply the wave function φ by $(-1)^l$, where l is the orbital angular momentum of the relative motion of the two particles (see §30). By comparing

---

[4] When there is degeneracy we can always choose linear combinations of the functions belonging to a given level, such that this condition is again satisfied.

these considerations with those given above, we conclude that a system of two identical particles with zero spin can have only an even orbital angular momentum.

Next, let the system consist of two particles with spin 1/2 (say, electrons). Then the complete wave function of the system (i.e. the product of the function $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ and the spin function $\chi(\sigma_1, \sigma_2)$) must certainly be antisymmetrical with respect to an interchange of the two particles. Hence, if the coordinate function is symmetrical, the spin function must be antisymmetrical, and vice versa. We shall write the spin function in spinor form, i.e. as a spinor $\chi^{\lambda\mu}$ of rank two, each of whose indices corresponds to the spin of one of the electrons. A symmetrical spinor ($\chi^{\lambda\mu} = \chi^{\mu\lambda}$) corresponds to a function symmetrical with respect to the spins of the two particles, and an antisymmetrical spinor ($\chi^{\lambda\mu} = -\chi^{\mu\lambda}$) to an antisymmetrical function. We know, however, that a symmetrical spinor of rank two describes a system with total spin unity, while an antisymmetrical spinor reduces to a scalar, corresponding to zero spin.

Thus we reach the following conclusion. The energy levels to which there correspond symmetrical solutions $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ of Schrödinger's equation can actually occur when the total spin of the system is zero, i.e. when the spins of the two electrons are "antiparallel", giving a sum of zero. The values of the energy belonging to antisymmetrical functions $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$, on the other hand, require a value of unity for the total spin, i.e. the spins of the two electrons must be "parallel".

In other words, the possible values of the energy of a system of electrons depend on their total spin. For this reason we can speak of a peculiar interaction of the particles which results in this dependence. This is called *exchange interaction*. It is a purely quantum effect, which entirely vanishes (like the spin itself) in the passage to the limit of classical mechanics.

The following situation is characteristic of the case of a system of two electrons which we have discussed. To each energy level there corresponds one definite value of the total spin, 0 or 1. This one-to-one correspondence between the spin values and the energy levels is preserved, as we shall see below (§63), in systems containing any number of electrons. It does not hold, however, for systems composed of particles whose spin exceeds 1/2.

Let us consider a system of two particles, each with arbitrary spin s. Its spin wave function is a spinor of rank $4s$:

$$\chi^{\overbrace{\lambda\mu\ldots}^{2s}\overbrace{\rho\sigma}^{2s}},$$

half ($2s$) of whose indices correspond to the spin of one particle, and the other half to that of the other particle. The spinor is symmetrical with respect to

the indices in each group. An interchange of the two particles corresponds to an interchange of all the indices $\lambda, \mu, \ldots$ of the first group with the indices $\rho, \sigma, \ldots$ of the second group. In order to obtain the spin function of a state of the system with total spin S, we must contract this spinor with respect to $2s - S$ pairs of indices (each pair containing one index from $\lambda, \mu, \ldots$ and one from $\rho, \sigma, \ldots$), and symmetrize it with respect to the remainder; as a result we obtain a symmetrical spinor of rank $2S$. However, the contraction of a spinor with respect to a pair of indices means, as we know, the construction of a combination antisymmetrical with respect to these indices. Hence, when the particles are interchanged, the spin wave function is multiplied by $(-1)^{2s-S}$.

On the other hand, the complete wave function of a system of two particles must be multiplied by $(-1)^{2s}$ when they are interchanged (i.e. by $+1$ for integral s and by -1 for half-integral s). Hence it follows that the symmetry of the coordinate wave function with respect to an interchange of the particles is given by the factor $(-1)^S$, which depends only on S. Thus we reach the result that the coordinate wave function of a system of two identical particles is symmetrical when the total spin is even, and antisymmetrical when it is odd.

Recalling what was said above concerning the relation between interchange of the particles and inversion of the coordinate system, we conclude also that, when the spin S is even (odd), the system can have only an even (odd) orbital angular momentum.

We see that here also a certain dependence is revealed between the possible values of the energy of the system and the total spin, but this dependence is not necessarily one-to-one. The energy levels to which there correspond symmetrical (antisymmetrical) coordinate wave functions can occur for any even (odd) value of S.

Let us calculate how many different states of the system there are with even and odd S. The quantity S takes $2s + 1$ values: $2s, 2s - 1, \ldots 0$. For any given S there are $2S + 1$ states differing in the value of the z-component of the spin ($(2s + 1)^2$ different states altogether). Let s be integral. Then, among the $2s + 1$ values of S, $s + 1$ are even and s odd. The total number of states with even S is equal to the sum

$$\sum_{S=0,2,\ldots,2s} (2S + 1) = (2s + 1)(s + 1);$$

the remaining $s(2s + 1)$ states have odd S. Similarly, we find that, when s is half-integral, there are $s(2s+1)$ states with even values of S and $(s+1)(2s+1)$ with odd values.

PROBLEMS

1. Determine the exchange splitting of the energy levels of a system of two electrons, regarding the interaction of the electrons as a perturbation.

SOLUTION. Let the particles be (when their interaction is neglected) in states with orbital wave functions $\varphi_1(\boldsymbol{r}_1)$ and $\varphi_2(\boldsymbol{r}_2)$. The states of the system with total spin $S = 0$ and $S = 1$ correspond to symmetrized and antisymmetrized products respectively:

$$\varphi = \frac{1}{\sqrt{2}}\left[\varphi_1(\boldsymbol{r}_1)\varphi_2(\boldsymbol{r}_2) \pm \varphi_1(\boldsymbol{r}_2)\varphi_2(\boldsymbol{r}_1)\right].$$

The mean value of the operator of the interaction $U(\boldsymbol{r}_2 - \boldsymbol{r}_1)$ of the particles in these states is $A \pm J$, where

$$A = \iint U|\varphi_1(\boldsymbol{r}_1)|^2|\varphi_2(\boldsymbol{r}_2)|^2\mathrm{d}V_1\mathrm{d}V_2,$$

$$J = \iint U\varphi_1(\boldsymbol{r}_1)\varphi_1^*(\boldsymbol{r}_2)\varphi_2(\boldsymbol{r}_2)\varphi_2^*(\boldsymbol{r}_1)\mathrm{d}V_1\mathrm{d}V_2$$

the latter being called the *exchange integral.* Omitting the additive constant A, which is not an exchange term, we therefore find the level shifts $\Delta E_0 = J, \Delta E_1 = -J$ (where the suffix indicates the value of S). These quantities can be represented as the eigenvalues of the spin *exchange operator*[5]

$$\widehat{V}_{\text{exch}} = -(1/2)J(1 + 4\widehat{\boldsymbol{s}}_1\widehat{\boldsymbol{s}}_2) \tag{1}$$

the eigenvalues of the product $\widehat{\boldsymbol{s}}_1\widehat{\boldsymbol{s}}_2$ are derived in §56, Problem 2.

If the electrons belong to different atoms, for example, the exchange integral decreases exponentially with increasing distance R between the atoms. It is clear from the form of the integrand that this integral is determined by the "overlap" of the wave functions of the states $\varphi_1(\boldsymbol{r}_1)$ and $\varphi_2(\boldsymbol{r}_2)$; using the asymptotic law of decrease of the wave functions of states of a discrete spectrum (cf. (21.6)), we find that

$$J \sim \exp(-(\varkappa_1 + \varkappa_2)R), \quad \varkappa_1 = \frac{1}{\hbar}\sqrt{2m|E|}, \quad \varkappa_2 = \frac{1}{\hbar}\sqrt{2m|E_2|},$$

where E_1 and E_2 are the energy levels of the electron in the two atoms.

2. The same as Problem 1, but for a system of three electrons.

SOLUTION. Using formula (1), Problem 1, we can write the operator of pairwise exchange interaction in a system of three electrons as

$$\widehat{V}_{\text{exch}} = -\sum J_{ab}(1/2 + 2\widehat{\boldsymbol{s}}_a\widehat{\boldsymbol{s}}_b) \tag{1}$$

where the summation is over pairs of particles 12, 13 and 23. The matrix elements of the operators $\widehat{\boldsymbol{s}}_a\widehat{\boldsymbol{s}}_b$ between states with different values of the pair of numbers σ_a, σ_b are given by formulae (55.6) as

$$\langle 1/2, 1/2|\boldsymbol{s}_a\boldsymbol{s}_b|1/2, 1/2\rangle = 1/4, \quad \langle 1/2, -1/2|\boldsymbol{s}_a\boldsymbol{s}_b|1/2, -1/2\rangle = -1/4,$$

---

[5] First used by Dirac.

$$\langle 1/2, -1/2|\boldsymbol{s}_a\boldsymbol{s}_b| - 1/2, 1/2\rangle = 1/2.$$

We first determine the energy corresponding to the greatest possible value of the total-spin component $M_s = \sigma_1 + \sigma_2 + \sigma_3$, viz. $M_s = 3/2$. This gives the energy of the state with total spin $S = 3/2$. On calculating the corresponding diagonal matrix element of the operator (1), we find

$$\Delta E_{3/2} = -(J_{12} + J_{13} + J_{23}).$$

Next we take states with $M_s = 1/2$ This value can occur in three ways, depending on which of the numbers $\sigma_1, \sigma_2, \sigma_3$ is $-1/2$ (the other two being $1/2$). Thus for these states we should have a secular equation of the third degree. The calculation can, however, be simplified immediately by noting that one of the roots of this equation must correspond to the energy already found for the state with $S = 3/2$, and the secular equation must therefore have the factor $\Delta E - \Delta E_{3/2}$. In this way the calculation of the free term in the cubic equation can be avoided.[6]

The leading terms of the equation are found to be

$$(\Delta E)^3 + (J_{12} + J_{13} + J_{23})(\Delta E)^2 +$$
$$+ [J_{12}J_{13} + J_{12}J_{23} + J_{13}J_{23} - (J_{12}^2 + J_{13}^2 + J_{23}^2)]\Delta E + \cdots = 0,$$

Dividing by $\Delta E + J_{12} + J_{13} + J_{23}$, we find the two energy levels corresponding to states with spins $S = 1/2$:

$$\Delta E_{1/2} = \pm[J_{12}^2 + J_{13}^2 + J_{23}^2 - J_{12}J_{13} - J_{12}J_{23} - J_{13}J_{23}]^{1/2}.$$

Thus there are three energy levels, in accordance with the calculation in §63, Problem 1

3. In which states can the ^{8}Be nucleus decay into two α-particles?

SOLUTION. Since the α-particle has no spin, a system of two -particles can only have an even orbital angular momentum (equal to the total angular momentum), and its states are even. The decay in question is therefore possible only from even states of the ^{8}Be nucleus with even total angular momentum.

§ 63. Symmetry with respect to interchange

By considering a system composed of only two particles, we have been able to show that its coordinate wave functions $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ for the stationary states must be either symmetrical or antisymmetrical. In the general case of a system of an arbitrary number of particles, the solutions of Schrödinger's

---

[6] This device is particularly useful in similar calculations for systems with a larger number of particles.

equation (the coordinate wave functions) need not necessarily be either symmetrical or antisymmetrical with respect to the interchange of any pair of particles, as the complete wave functions (which include the spin factor) must be. This is because an interchange of only the coordinates of two particles does not correspond to a physical interchange of them. The physical identity of the particles here leads only to the fact that the Hamiltonian of the system is invariant with respect to the interchange of the particles, and hence, if some function is a solution of Schrödinger's equation, the functions obtained from it by various interchanges of the variables will also be solutions.

Let us first of all make some remarks regarding interchanges in general. In a system of N particles, $N!$ different permutations in all are possible. If we imagine all the particles to be numbered, each permutation can be represented by a definite sequence of the numbers $1, 2, 3, \ldots$. Every such sequence can be obtained from the natural sequence $1, 2, 3, \ldots$ by successive interchanges of pairs of particles. The permutation is called even or odd, according as it is brought about by an *even* or *odd* number of such interchanges. We denote by the $\widehat{P}$ operators of permutations of N particles, and introduce a quantity δP which is $+1$ if is an even permutation and -1 if it is odd. If φ is a function symmetrical with respect to all the particles, we have

$$\widehat{P}\varphi = \varphi,$$

while, if φ is antisymmetrical with respect to all the particles, then

$$\widehat{P}\varphi = \delta_P \varphi.$$

From an arbitrary function $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$, we can form a symmetrical function by the operation of *symmetrization*, which can be written

$$\varphi_{\text{sym}} = \text{const} \sum_P \widehat{P}\varphi, \tag{63.1}$$

where the summation extends over all possible permutations. The formation of an antisymmetrical function (an operation sometimes called *alternation*) can be written as

$$\varphi_{\text{ant}} = \text{const} \sum_P \delta_P \widehat{P}_\varphi. \tag{63.2}$$

Let us return to considering the behaviour, with respect to permutations, of the wave functions φ of a system of identical particles.[7] The fact that the

---

[7] From the mathematical point of view, the problem is to find irreducible representations of the permutation group. A detailed account of the mathematical theory of permutation (or symmetry) groups is given by H. Weyl, *The Theory of Groups and Quantum Mechanics*, Methuen, London 1931; M. Hamermesh, *Group Theory and its Application to Physical Problems*, Pergamon, London, 1962; I. G. Kaplan, *Symmetry of Many-Electron Systems*, Academic Press, New York, 1974.

Hamiltonian $\widehat{H}$ of the system is symmetrical with respect to all the particles means, mathematically, that $\widehat{H}$ commutes with all the permutation operators $\widehat{P}$. These operators, however, do not commute with one another, and so they cannot be simultaneously brought into diagonal form. This means that the wave functions φ cannot be so chosen that each of them is either symmetrical or antisymmetrical with respect to all interchanges separately.[8])

Let us try to determine the possible types of symmetry of the functions $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$ of N variables (or of sets of several such functions) with respect to permutations of the variables. The symmetry must be such that it cannot be increased, i.e. such that any additional operation of symmetrization or alternation, on being applied to these functions, would reduce them either to linear combinations of themselves or to zero identically.

We already know two operations which give functions with the greatest possible symmetry: symmetrization with respect to all the variables, and alternation with respect to all the variables. These operations can be generalized as follows.

We divide the set of all the N variables $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ (or, what is the same thing, the suffixes $1, 2, 3, \ldots, N$) into several sets, containing $N_1, N_2, \ldots$ elements (variables); $N_1 + N_2 + \cdots = N$. This division can be conveniently shown by a diagram (known as a *Young diagram*) in which each of the numbers $N_1, N_2, \ldots$ is represented by a line of several cells (thus, Fig. 21 gives a diagram of the divisions $6 + 4 + 4 + 3 + 3 + 1 + 1$ and $7 + 5 + 5 + 3 + 1 + 1$ for $N = 22$); one of the numbers $1, 2, 3, \ldots$ is to be placed in each square. If we place the lines in order of decreasing length (as in Fig. 10), the diagram contains not only successive horizontal rows, but also vertical columns.

Let us symmetrize an arbitrary function $\varphi(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_N)$ with respect to the variables in each row. The alternation operation can then be performed only with respect to the variables in different rows; alternation with respect to a pair of variables in the same row clearly gives zero identically.

Having chosen one variable from each row, we can, without loss of generality, regard them as being in the first cells in each row (after symmetrization, the order of the variables among the cells in each row is immaterial); let us alternate with respect to these variables. Having then deleted the first column, we alternate with

FIG.21

[8]) Except for a system of only two particles, where there is a single interchange operator, which can be brought into diagonal form simultaneously with the Hamiltonian.

respect to variables chosen one from each row in the thus "curtailed" diagram; these variables can again be regarded as being in the first cells of the "curtailed" rows. Continuing this process, we finally have the function first symmetrized with respect to the variables in each row and then alternated with respect to the variables in each column. After alternation, of course, the function in general ceases to be symmetrical with respect to the variables in each row. The symmetry is preserved only with respect to the variables in the cells of the first row which project beyond the other rows.

Having distributed the N variables in various ways among the rows of a Young diagram (the distribution among the cells in each row is immaterial), we thus obtain a series of functions, which are transformed linearly into one another when the variables are permuted in any manner.[9] However, it must be emphasized that not all these functions are linearly independent; the number of independent functions is in general less than the number of possible distributions of the variables among the rows of the diagram. We shall not pause here, however, to discuss this more closely.[10]

Thus any Young diagram determines some type of symmetry of functions with respect to permutations. By constructing all the possible Young diagrams (for a given N), we find all possible types of symmetry. This amounts to dividing the number N in all possible ways into a sum of smaller terms, including the number N itself; thus for $N = 4$ the possible partitions are $4, 3 + 1, 2 + 2, 2 + 1 + 1, 1 + 1 + 1 + 1$.

To each energy level of the system we can make correspond a Young diagram which determines the permutational symmetry of the appropriate solutions of Schrödinger's equation; in general, several different functions correspond to each value of the energy, and these are transformed linearly into each other by permutations. The existence of this "permutational degeneracy" is related to the fact that the operators $\widehat{P}$ each commute with the Hamiltonian but not with one another (see the middle of §10). However, it must be emphasized that this does not signify any additional physical degeneracy of the energy levels. All these different coordinate wave functions, multiplied by the spin functions, enter into a single definite combination—the complete wave function—which satisfies (according to the spin of the

---

[9] It would be possible to perform the symmetrization and alternation in the reverse order: to alternate with respect to the variables in each column, and then to symmetrize with respect to those in the rows. This, however, would give effectively the same thing, since the functions obtained by the two methods are linear combinations of one another.

[10] The independent functions that are transformed into linear combinations of one another form the basis of an irreducible representation of the permutation group. Their number is the dimension of the representation. For particles with spin 1/2 the number is derived in Problem 1 below.

particles) the condition of symmetry or antisymmetry.

Among the various types of symmetry there are always (for any given N) two to each of which only one function corresponds. One of these corresponds to a function symmetrical with respect to all the variables, and the other to one which is similarly antisymmetrical; in the first case, the Young diagram consists of a single row of N cells, and in the second case of a single column.

Let us now consider the spin wave functions $\chi(\sigma_1, \sigma_2, \ldots, \sigma_N)$. Their kinds of symmetry with respect to permutations of the particles are given by the same Young diagrams, with the components of the spins of the particles taking the part of variables. There arises the question of what diagram must correspond to the spin function for a given diagram of the coordinate function. Let us first suppose that the spin of the particles is integral. Then the complete wave function ψ must be symmetrical with respect to all the particles. For this to be so, the symmetry of the spin and coordinate functions must be given by the same Young diagram, and the complete wave function ψ is expressed as definite bilinear combinations of the two; we shall not here pause to examine more closely the problem of constructing these combinations.

Next, suppose the spin of the particles to be half-integral. Then the complete wave function must be antisymmetrical with respect to all the particles. It can be shown that, for this to be so, the Young diagrams for the coordinate and spin functions must be in dual relation, i.e. obtained from each other by interchanging rows and columns (as in the two diagrams shown in Fig. 10)

Let us consider in more detail the important case of particles with spin $1/2$ (electrons, for instance). Each of the spin variables $\sigma_1, \sigma_2, \ldots$ here takes only the two values $\pm 1/2$. Since a function antisymmetrical with respect to any two variables vanishes when these variables take the same value, it is clear that the function χ can be alternated only with respect to pairs of variables; if we alternate with respect to even three variables, two of them must always take the same value, so that we have zero identically.

Thus, for a system of electrons, the Young diagrams for the spin functions can contain columns of only one or two cells (i.e. only one or two rows); in the Young diagrams for the coordinate functions, the same is true of the number of columns. The number of possible types of permutational symmetry for a system of N electrons is therefore equal to the number of possible partitions of the number N into a sum of ones and twos. When N is even, this number is $N/2+1$ (partitions with $0, 1, \ldots, N/2$ twos), while if N is odd it is $(N+1)/2$ (partitions with $0, 1, \ldots, (N-1)/2$ twos). Thus, for instance, Fig. 11 shows the possible Young diagrams (coordinate and spin) for $N = 4$.

It is easy to see that each of these types of symmetry (i.e. each of the Young diagrams) corresponds to a definite total spin S of the system of

electrons. We shall consider the spin functions in spinor form, i.e. as spinors $\chi^{\lambda\mu\cdots}$ of rank N, whose indices (each of which corresponds to the spin of an individual particle) will be the variables that are arranged in the cells of the Young diagrams. Let us examine the Young diagram consisting of two rows with N_1 and N_2 cells ($N_1 + N_2 = N$, and $N_1 \geqslant N_2$). In each of the first N_2 columns there are two cells, and the spinor must be antisymmetrical with respect to the corresponding pairs of indices. With respect to the indices in the last $n = N_1 - N_2$ cells in the first row, however, it must be symmetrical. As we
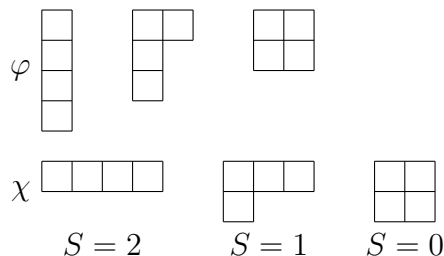
FIG.22

know, such a spinor of rank N reduces to a symmetrical spinor of rank n, to which there corresponds a total spin $S = n/2$. Returning to the Young diagrams for the coordinate functions, we can say that the diagram with n rows each of one cell corresponds to a total spin $S = n/2$. For even N, the total spin can take integral values from 0 to $N/2$, while for odd N it can take half-integral values from $1/2$ to $N/2$, as it should.

We emphasize that this one-to-one correspondence between the Young diagrams and the total spin holds only for systems of particles with spin $1/2$; we have seen this, for a system of two particles, in the previous section. For a system of N particles with spin s, the spin wave function is made up of a product of N symmetrical spinors of rank $2s$, i.e. is a spinor of rank $2Ns$. If this spinor is symmetrized according to a particular Young diagram of N cells, we can usually construct from the independent components of the symmetrized spinor several sets of linear combinations, each set corresponding to a different total spin S of the system.

In the same way as the Young diagram for the spin functions of particles with spin $1/2$ cannot contain columns of more than two cells, so for particles with any spin s the columns cannot contain more than $2s + 1$ cells.

If the number N of particles in the system is an integral multiple of $2s + 1$, the possible Young diagrams include a rectangle with $2s + 1$ cells in each column. This corresponds to one definite value of the total spin, $S = 0$. Hence we can conclude that the same value of S corresponds to any two (spin) Young diagrams which can be fitted together to form a rectangle of

height $2s + 1$.[11]) This is a simple consequence of the fact that the addition of two angular momenta can give zero only if they have the same absolute magnitude.

To conclude this section, let us return to the fact already mentioned in the footnote at the end of §20 that, for a system of several identical particles, we cannot assert that the wave function of the stationary state of lowest energy is without nodes. We can now amplify this statement and elucidate its origin.

The wave function (that is, the coordinate function), if it has no nodes, must certainly be symmetrical with respect to all the particles; for, if it were antisymmetrical with respect to the interchange of any pair of particles $1, 2$, it would vanish for $\boldsymbol{r}_1 = \boldsymbol{r}_2$. If, however, the system consists of three or more electrons, no completely symmetrical coordinate wave function is possible (the Young diagram of the coordinate function cannot have rows with more than two cells). Thus, although the solution of Schrödinger's equation which corresponds to the lowest eigenvalue is without nodes (by the theorem of the variational calculus), this solution may be physically inadmissible; the smallest eigenvalue of Schrödinger's equation will not then correspond to the normal state of the system, and the wave function of this state will in general have nodes. For particles with a half-integral spin s, this situation occurs in systems with more than $2s + 1$ particles. For systems of bosons, a completely symmetrical coordinate wave function is always possible.

PROBLEMS

1. Determine the number of energy levels with different values of the total spin S, for a system of N particles with spin $1/2$.

SOLUTION. A given value of the projection of the total spin of the system, $M_S = \sum \sigma$, can be obtained in

$$f(M_S) = \frac{N!}{\left(\frac{N}{2} + M_S\right)! \left(\frac{N}{2} - M_S\right)!}$$

ways, with $N/2 + M_S$ particles taken to have $\sigma = 1/2$ and the remainder $\sigma = -1/2$. To each energy level with a given S, there correspond $2S + 1$ states with $M_S = S, S - 1, \ldots, -S$. Hence it is easy to see that the number of different energy levels with a given value of S is

$$n(S) = f(S) - f(S + 1) = \frac{N!(2S + 1)}{\left(\frac{N}{2} + S + 1\right)! \left(\frac{N}{2} - S\right)!}.$$

---

[11]) For example, the two diagrams (for $s = 1$):
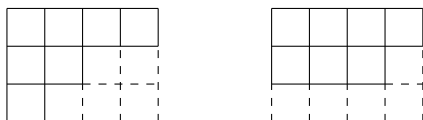
The continuous and broken lines show the complementary diagrams.

The total number of different energy levels is

$$n = \sum_S n(S) = f(0) = \frac{N!}{\left(\frac{N}{2}!\right)^2}$$

for even N, and

$$n = f\left(\frac{1}{2}\right) = \frac{N!}{\left(\frac{N+1}{2}\right)!\left(\frac{N-1}{2}\right)!}$$

for odd N.

2. Find the values of the total spin S that occur for various types of symmetry of the spin functions of a system of two, three or four particles with spin 1.

SOLUTION. For two particles, the correspondence is established by the fact that the factor by which the spin function is multiplied when the particles are interchanged must be $(-1)^{2s-S}$ (see the end of §62). For particles with spin $s = 1$ this gives

$$(1)$$

The Young diagrams for a system of three particles are obtained by adding to the diagrams (1) one cell in every possible way. The result may be written as the symbolic equations

The values of S are shown beneath each diagram, and the values of the total spin of the system of three particles (the diagrams on the right) are found from the spins of the two-particle and one-particle systems (the diagrams on the left) by the rule of addition of angular momenta.[12] The distribution of the resulting values of S among the diagrams on the right is established by noting that diagram (c) (a column of three cells) corresponds to $S = 0$, and (b) therefore to the remaining

---

[12] The repetition of 1 beneath the right-hand diagrams occurs because this value of the angular momentum comes firstly from adding the angular momenta 0 and 1, and secondly from adding 2 and 1.

values 1 and 2 in the second equation, while (a) belongs to the values 1 and 3 that are left after (b) has been labelled in the first equation:

$$(2)$$

The Young diagrams for a system of four particles are obtained by adding one cell to the diagrams (2), with the condition that no column should contain more than three cells:

Diagram (c) can be added to (a) in (1) to form a rectangle with three-cell columns, and therefore corresponds to the same values $S = 0, 2$. The values of S for diagram (b) are found from the remainder of the second equation, and then those for (a) from the remainder of the first equation:

§ 64. Second quantization. The case of Bose statistics

In the theory of systems consisting of a large number of identical particles, there is a widely used method of considering the problem, known as

second quantization. This method is especially necessary in relativistic theory, where we have to deal with systems in which the number of particles is itself variable.[13]

Let $\psi_1(\xi), \psi_2(\xi), \ldots$ be some complete set of orthogonal and normalized wave functions of stationary states of a single particle.[14] These may be states of a particle in some arbitrarily chosen external field, but are usually taken to be simply plane waves, i.e. the wave functions of a free-particle having definite values of the momentum (and spin projection). In order to make the spectrum of states discrete, we shall consider the motion of particles in a large but finite region, for which the eigenvalues of the momentum components form a discrete series, the intervals between adjacent values being inversely proportional to the linear dimensions of the region and tending to zero as these increase.

In a system of free particles, the particle momenta are separately conserved. The occupation numbers of the states are therefore also conserved, i.e. the numbers $N_1, N_2, \ldots$ which show how many particles are in each of the states $\psi_1, \psi_2, \ldots$. In a system of interacting particles, the momentum of each particle is not conserved, and so the occupation numbers are not conserved. For such a system we can consider only the probability distribution of the various values of the occupation numbers. Let us seek to construct a mathematical formalism in which the occupation numbers (and not the coordinates and spin projections of the particles) play the part of independent variables.

In this formalism, it is convenient to use the Dirac notation (see the end of §11), taking $N_1, N_2, \ldots$ as quantum numbers defining the state. The states corresponding to the wave functions (61.3) and (61.5) will be denoted by $|N_1, N_2, \ldots\rangle$. The coordinate and spin variables of the particles are not shown explicitly.

In accordance with this choice of the independent variables, the operators of the various physical quantities (including the Hamiltonian of the system) must be formulated in terms of their action on functions of the occupation numbers. Such a formulation can be obtained on the basis of the usual matrix representation of operators. The operator matrix elements must be considered in relation to the wave functions of the stationary states of a system of non-interacting particles. Since these states can be described by

---

[13] The method of second quantization was developed by P. A. M. Dirac (1927) for photons in radiation theory, and later extended to fermions by E. Wigner and P. Jordan (1928).

[14] As in §61, ξ denotes the set of the coordinates and the spin projection σ of the particle, and integration with respect to ξ is taken to mean integration over the coordinates and summation over σ.

specifying definite values of the occupation numbers, this will also show the nature of the action of the operators on these variables.

Let us first consider systems of particles obeying Bose statistics. Let $\widehat{f}_a^{(1)}$ be the operator of some quantity pertaining to the ath particle, i.e. acting only on functions of the variables ξ_a. We introduce the operator

$$\widehat{F}^{(1)} = \sum_a \widehat{f}_a^{(1)} \tag{64.1}$$

which is symmetrical with respect to all the particles (the summation being over all particles), and determine its matrix elements with respect to the wave functions (61.3). First of all, it is easy to see that the matrix elements will be different from zero only for transitions which leave the numbers $N_1, N_2, \ldots$ unchanged (diagonal elements) and for transitions where one of these numbers is increased, and another decreased, by unity. For, since each of the operators $\widehat{f}_a^{(1)}$ acts only on one function in the product $\psi_{p_1}(\xi_1)\psi_{p_2}(\xi_2)\ldots\psi_{p_N}(\xi_N)$, its matrix elements can be different from zero only for transitions whereby the state of a single particle is changed; this, however, means that the number of particles in one state is diminished by unity, while the number in another state is correspondingly increased. The calculation of these matrix elements is in principle very simple; it is easier to do it oneself than to follow an account of it. Hence we shall give only the result of this calculation. The non-diagonal elements are

$$\left\langle N_i, N_k - 1 | F^{(1)} | N_i - 1, N_k \right\rangle = f_{ik}^{(1)} \sqrt{N_i N_k}. \tag{64.2}$$

We shall indicate only those suffixes with respect to which the matrix element is non-diagonal, omitting the remainder for brevity. Here $f_{ik}^{(1)}$ is the matrix element

$$f_{ik}^{(1)} = \int \psi_i^*(\xi)\widehat{f}^{(1)}\psi_k(\xi)\mathrm{d}\xi; \tag{64.3}$$

since the operators $f_a^{(1)}$ differ only in the naming of the variables on which they act, the integrals (64.3) are independent of a, which is therefore omitted. The diagonal matrix elements of $F^{(1)}$ are the mean values of the quantity $F^{(1)}$ in the states $\Psi_{N_1 N_2 \ldots}$. Calculation gives

$$\overline{F^{(1)}} = \sum_i f_{ii}^{(1)} N_i. \tag{64.4}$$

We now introduce the operators $\widehat{a}_i$, which play a leading part in the method of second quantization; they act, not on functions of the coordinates, but on functions of the occupation numbers. By definition, the operator $\widehat{a}_i$

acting on the function $|N1, N2, \ldots\rangle$ decreases the value of the variable N_i by unity, and at the same time it multiplies the function by $\sqrt{N_i}$:[15]

$$\widehat{a}_i \,|N_1, N_2, \ldots, N_i, \ldots\rangle = \sqrt{N_i}\,|N_1, N_2, \ldots, N_i - 1, \ldots\rangle. \qquad (64.5)$$

We can say that the operator $\widehat{a}_i$ diminishes by one the number of particles in the ith state; it is therefore called a particle *annihilation operator*. It can be represented in the form of a matrix whose only non-zero element is

$$\langle N_i - 1|a_i|N_i\rangle = \sqrt{N_i}. \qquad (64.6)$$

The operator $\widehat{a}_i^\dagger$ which is the Hermitian conjugate of $\widehat{a}_i$ is, by definition (see (11.9)), represented by a matrix whose only non-zero element is

$$\langle N_i|\, a_i^\dagger\,|N_i - 1\rangle = \langle N_i - 1|a_i|N_i\rangle^* = \sqrt{N_i} \qquad (64.7)$$

This means that, when acting on the function $\Phi(N_1, N_2, \ldots)$. it increases the number N_i by unity:

$$\widehat{a}_i^\dagger \,|N_1, N_2, \ldots, N_i, \ldots\rangle = \sqrt{N_i + 1}\,|N_1, N_2, \ldots, N_i + 1, \ldots\rangle. \qquad (64.8)$$

In other words, the operator $\widehat{a}_i^\dagger$ increases by one the number of particles in the ith state, and is therefore called a particle *creation operator*.

The product of the operators $\widehat{a}_i^\dagger \widehat{a}_i$, acting on the wave function, must multiply it by a constant simply, leaving unchanged all the variables $N_1, N_2, \ldots$: the operator $\widehat{a}_i$ diminishes N_i by unity, and $\widehat{a}_i^\dagger$ then restores it to its original value. Direct multiplication of the matrices (64.6) and (64.7) shows that $\widehat{a}_i^\dagger \widehat{a}_i$ is represented, as we should expect, by a diagonal matrix whose diagonal elements are N_i. We can write

$$\widehat{a}_i^\dagger \widehat{a}_i = N_i. \qquad (64.9)$$

Similarly, we find that

$$\widehat{a}_i \widehat{a}_i^\dagger = N_i + 1. \qquad (64.10)$$

The difference of these equations gives the commutation rule for the operators $\widehat{a}_i$ and $\widehat{a}_i^\dagger$:

$$\widehat{a}_i \widehat{a}_i^\dagger - \widehat{a}_i^\dagger \widehat{a}_i = 1. \qquad (64.11)$$

The operators with i and k different act on different variables (N_i and N_k), and commute:

$$\widehat{a}_i \widehat{a}_k - \widehat{a}_k \widehat{a}_i = 0, \quad \widehat{a}_i \widehat{a}_k^\dagger - \widehat{a}_k^\dagger \widehat{a}_i = 0, \quad i \neq k. \qquad (64.12)$$

---

[15] Here we use the notation $\widehat{a}\,|m\rangle$ with the natural sense of the result of the operator $\widehat{a}$ acting on the state wave function $|m\rangle$.

From the above properties of the operators $\widehat{a}_i, \widehat{a}_i^\dagger$ it is easy to see that the operator

$$\widehat{F}^{(1)} = \sum_{i,k} = f_{ik}^{(1)}\widehat{a}_i^\dagger\widehat{a}_k \tag{64.13}$$

is the same as the operator (64.1). For all the matrix elements calculated from (64.6), (64.7) are the same as the elements (64.2), (64.4). This is a very important result. In formula (64.13), the quantities $f_{ik}^{(1)}$ are simply numbers. Thus we have been able to express an ordinary operator, acting on functions of the coordinates, in the form of an operator acting on functions of new variables, the occupation numbers N_i.

The result which we have obtained is easily generalized to operators of other forms. Let

$$\widehat{F}^{(2)} = \sum_{a>b} \widehat{f}_{ab}^{(2)}, \tag{64.14}$$

where $\widehat{f}_{ab}^{(2)}$ is the operator of a physical quantity pertaining to two particles at once, and hence acts on functions of ξ_a and ξ_b. Similar calculations show that this operator can be expressed in terms of the operators $\widehat{a}_i, \widehat{a}_i^\dagger$ by

$$\widehat{F}^{(2)} = \frac{1}{2} \sum_{i,k,l,m} \langle ik|f^{(2)}|lm\rangle \widehat{a}_i^\dagger\widehat{a}_k^\dagger\widehat{a}_m\widehat{a}_l, \tag{64.15}$$

where

$$\langle ik|f^{(2)}|lm\rangle = \iint \psi_i^*(\xi_1)\psi_k^*(\xi_2)\widehat{f}^{(2)}\psi_l(\xi_1)\psi_m(\xi_2)\mathrm{d}\xi_1\mathrm{d}\xi_2.$$

The generalization of these formulae to operators of any other form symmetrical with respect to all the particles (of the form $\widehat{F}^{(3)} = \sum \widehat{f}_{abc}^{(3)}$, etc.) is obvious.

These formulae can be used to express, in terms of the operators $\widehat{a}_i$ and $\widehat{a}_i^\dagger$, the Hamiltonian of the physical system of N identical interacting particles that is being considered. The Hamiltonian of such a system is, of course, symmetrical with respect to all the particles. In the non-relativistic approximation,[16] it is independent of the spins of the particles, and can be represented in a general form as follows:

$$\widehat{H} = \sum_a \widehat{H}_a^{(1)} + \sum_{a>b} U^{(2)}(\boldsymbol{r}_a, \boldsymbol{r}_b) + \sum_{a>b>c} U^{(3)}(\boldsymbol{r}_a, \boldsymbol{r}_b, \boldsymbol{r}_c) + \dots \tag{64.16}$$

Here $\widehat{H}_a^{(1)}$ is the part of the Hamiltonian which depends on the coordinates of the ath particle only:

$$\widehat{H}_a^{(1)} = -\frac{\hbar^2}{2m}\Delta_a + U^{(1)}(\boldsymbol{r}_a). \tag{64.17}$$

---

[16] n the absence of a magnetic field.

where $U^{(1)}(\boldsymbol{r}_a)$ is the potential energy of a single particle in the external field. The remaining terms in (64.16) correspond to the mutual interaction energy of the particles; the terms depending on the coordinates of two, three, etc. particles have been separated.

This representation of the Hamiltonian enables us to apply formulae (64.13), (64.15) and their analogues directly. Thus

$$\widehat{H} = \sum_{i,k} H_{ik}^{(1)}\widehat{a}_i^\dagger\widehat{a}_k + \frac{1}{2}\sum_{i,k,l,m}\left\langle ik|U^{(2)}|lm\right\rangle\widehat{a}_i^\dagger\widehat{a}_k^\dagger\widehat{a}_m\widehat{a}_l\ldots \qquad (64.18)$$

This gives the required expression for the Hamiltonian in the form of an operator acting on functions of the occupation numbers.

For a system of non-interacting particles, only the first term in the expression (64.18) remains:

$$\widehat{H} = \sum_{i,k} H_{ik}^{(1)}\widehat{a}_i^\dagger\widehat{a}_k. \qquad (64.19)$$

If the functions ψ_i are taken to be the eigenfunctions of the Hamiltonian $\widehat{H}^{(1)}$ of an individual particle, the matrix $\widehat{H}_{ik}^{(1)}$ is diagonal, and its diagonal elements are the eigenvalues ε_i of the energy of the particle. Thus,

$$\widehat{H} = \sum_i \varepsilon_i \widehat{a}_i^\dagger\widehat{a}_i;$$

replacing the operator $\widehat{a}_i^\dagger\widehat{a}_i$ by its eigenvalues (64.9), we have for the energy levels of the system the expression

$$E = \sum_i \varepsilon_i N_i$$

a trivial result which could have been foreseen.

The formalism which we have developed can be put in a more compact form by introducing the ψ-*operators*[17])

$$\widehat{\psi}(\xi) = \sum_i \psi_i(\xi)\widehat{a}_i \quad \widehat{\psi}^\dagger(\xi) = \sum_i \psi^*(\xi)\widehat{a}_i^\dagger, \qquad (64.20)$$

---

[17]) Note the analogy between (64.20) and the expansion

$$\psi = \sum a_i\psi_i$$

of any wave function in terms of a complete set of functions. Here it is "re-quantized", and this is the reason for the term *second quantization method*.

where the variables ξ are regarded as parameters. By what has been said above concerning the operators $\widehat{a}_i, \widehat{a}_i^\dagger$, it is clear that the operator $\widehat{\psi}$ decreases the total number of particles in the system by one, while $\widehat{\psi}^\dagger$ increases it by one.

It is easy to see that the operator $\widehat{\psi}^\dagger(\xi_0)$ creates a particle at the point ξ_0. For the result of the action of the operator $\widehat{a}_i^\dagger$ is to create a particle in a state with wave function $\psi_i(\xi)$. Hence it follows that the result of the action of the operator $\widehat{\psi}^\dagger(\xi_0)$ is to create a particle in a state with wave function

$$\sum_i \psi_i^*(\xi)\psi_i(\xi_0) = \delta(\xi - \xi_0),$$

which corresponds to a particle with definite values of the coordinates (and spin). Here we have used formula (5.12).[18]

The commutation rules for the ψ operators are obtained at once from those for the operators $\widehat{a}_i, \widehat{a}_i^\dagger$:

$$\widehat{\psi}(\xi)\widehat{\psi}(\xi') - \widehat{\psi}(\xi')\widehat{\psi}(\xi) = 0, \tag{64.21}$$

$$\widehat{\psi}(\xi)\widehat{\psi}^\dagger(\xi') - \widehat{\psi}^\dagger(\xi')\widehat{\psi}(\xi) = \sum_i \psi_i(\xi)\psi_i^*(\xi') = \delta(\xi - \xi'). \tag{64.22}$$

The second-quantized operator $\widehat{F}^{(1)}$ can be written by means of the ψ operators in the form

$$\widehat{F}^{(1)} = \int \widehat{\psi}^\dagger(\xi)\widehat{f}^{(1)}\widehat{\psi}(\xi)\mathrm{d}\xi \tag{64.23}$$

where it is understood that the operator $\widehat{f}^{(1)}$ acts on functions of the parameters ξ in $\widehat{\psi}(\xi)$. For, substituting $\widehat{\psi}$ and $\widehat{\psi}^\dagger$ in the form (64.20) and using the definition (64.3), we return to (64.13). Similarly, (64.15) becomes

$$\widehat{F}^{(2)} = \frac{1}{2}\iint \widehat{\psi}^\dagger(\xi)\widehat{\psi}^\dagger(\xi')\widehat{f}^{(2)}\widehat{\psi}(\xi')\widehat{\psi}(\xi)\mathrm{d}\xi\mathrm{d}\xi'. \tag{64.24}$$

In particular, the Hamiltonian of the system, expressed in terms of the ψ operators, is

$$\widehat{H} = \int \left\{ -\frac{\hbar^2}{2m}\widehat{\psi}^\dagger(\xi)\Delta\widehat{\psi}(\xi) + \widehat{\psi}^\dagger(\xi)U^{(1)}(\xi)\widehat{\psi}(\xi) \right\}\mathrm{d}\xi +$$

$$+ \frac{1}{2}\iint \widehat{\psi}^\dagger(\xi)\widehat{\psi}^\dagger(\xi')U^{(2)}(\xi,\xi')\widehat{\psi}(\xi')\widehat{\psi}(\xi)\mathrm{d}\xi\mathrm{d}\xi' + \ldots \tag{64.25}$$

---

[18] $\delta(\xi - \xi_0)$ conventionally denotes the product

$$\delta(x - x_0)\delta(y - y_0)\delta(z - z_0)\delta_{\sigma\sigma_0}.$$

The operator $\widehat{\psi}^\dagger(\xi)\widehat{\psi}(\xi)$, constructed from the ψ operators by analogy with the product $\psi^*\psi$ which determines the probability density for a particle in a state with wave function ψ, is called the *particle density operator*. The integral

$$\widehat{N} = \int \widehat{\psi}^\dagger\widehat{\psi}\mathrm{d}\xi \qquad (64.26)$$

represents in the second-quantization formalism the operator of the total number of particles in the system. For, substituting the ψ operators in the form (64.20) and using the normalization and the orthogonality of the wave functions, we have

$$\widehat{N} = \sum \widehat{a}_i^\dagger\widehat{a}_i.$$

Each term in this sum is the operator of the number of particles in the ith state; according to (64.9), its eigenvalues are equal to the occupation numbers N_i, and the sum of all these numbers is the total number of particles in the system.[19])

Lastly, if the system consists of bosons of various kinds, operators $\widehat{a}$ and $\widehat{a}^\dagger$ for each kind of particle must be defined in the second quantization method. It is evident that operators pertaining to particles of different kinds commute.

§ 65. Second quantization. The case of Fermi statistics

The basic theory of the method of second quantization remains wholly unchanged for systems of identical fermions, but the actual formulae for the matrix elements of quantities and for the operators $\widehat{a}_i$ are naturally different.

The wave function $\psi_{N_1 N_2\ldots}$ now has the form (61.5). Because of the antisymmetry of this function, the question of its sign arises first of all. This question did not arise in the case of Bose statistics, since, because of the symmetry of the wave function, its sign, once chosen, was preserved under all permutations of the particles. In order to make definite the sign of the function (61.5), we shall agree to choose it as follows. We number successively, once and for all, all the states ψ_i. We then complete the rows of the determinant (61.5) so that always

$$p_1 < p_2 < p_3 < \cdots < p_N, \qquad (65.1)$$

whilst in the successive columns we have functions of the different variables in the order $\xi_1, \xi_2, \ldots, \xi_N$. No two of the numbers $p_1, p_2, \ldots$ can be equal, since

---

[19]) For systems containing a specified number of particles these statements are trivial, as are the properties of the Hamiltonian (64.19) of a system of free particles. Their generalization in the relativistic theory, however, yields new results that are by no means trivial (cf. RQT, §11)

otherwise the determinant would vanish. In other words, the occupation numbers N_i can take only the values 0 and 1.

Let us again consider an operator of the form (64.1), $\widehat{F}^{(1)} = \sum \widehat{f}_a^{(1)}$. As in §64, its matrix elements will be non-zero only for transitions where all the occupation numbers remain unchanged and for those where one occupation number (N_i) is diminished by unity (becoming zero instead of one) and another (N_k) is increased by unity (becoming one instead of zero). We easily find that, for $i < k$,

$$\left\langle 1_i, 0_k | F^{(1)} | 0_i, 1_k \right\rangle = f_{ik}^{(1)} (-1)^{\sum(i+1, k-1)}. \tag{65.2}$$

where by $0_i, 1_i$ we signify $N_i = 0$, $N_i = 1$ and the symbol $\sum(k, l)$ denotes the sum of the occupation numbers of all states from the kth to the lth:[20]

$$\sum(k, l) = \sum_{n=k}^{l} N_n.$$

For the diagonal elements we obtain our previous formula (64.4):

$$\overline{F^{(1)}} = \sum_i f_{ii}^{(1)} N_i \tag{65.3}$$

In order to represent the operator $\widehat{F}^{(1)}$ in the form (64.13), the operators $\widehat{a}_i$ must be defined as matrices whose elements are

$$\langle 0_i | a_i | 1_i \rangle = \langle 1_i | a_i^\dagger | 0_i \rangle = (-1)^{\sum(1, i-1)}. \tag{65.4}$$

On multiplying these matrices, we find, for $k > i$,

$$\langle 1_i, 0_k | a_i^\dagger a_k | 0_i, 1_k \rangle = \langle 1_i, 0_k | a_i^\dagger | 0_i, 0_k \rangle \langle 0_i, 0_k | a_k | 0_i, 1_k \rangle =$$
$$= (-1)^{\sum(1, i-1)} (-1)^{\sum(1, i-1) + \sum(i+1, k-1)}$$

or

$$\langle 1_i, 0_k | a_i^\dagger a_k | 0_i, 1_k \rangle = (-1)^{\sum(i+1, k-1)}. \tag{65.5}$$

If $i = k$, the matrix of $\widehat{a}_i^\dagger \widehat{a}_i$ is diagonal, and its elements are unity for $N_i = 1$, and zero for $N_i = 0$; this can be written

$$\widehat{a}_i^\dagger \widehat{a}_i = N_i. \tag{65.6}$$

On substituting these expressions in (64.13), we in fact obtain (65.2), (65.3).

---

[20] For $i > k$ the exponent in (65.2) becomes $\sum(k+1, i-1)$. The sum must be taken as zero when $i = k \pm 1$.

Multiplying $\widehat{a}_i^\dagger, \widehat{a}_k$ in the opposite order, we have

$$\langle 1_i, 0_i| \, a_k \widehat{a}_i^\dagger \, |0_i, 1_k\rangle = \langle 1_i, 0_k|a_k|1_i, 1_k\rangle \, \langle 1_i, 1_k| \, a_i^\dagger \, |0_i, 1_k\rangle =$$
$$= (-1)^{\sum(1,i-1)+\sum(i+1,k-1)+\sum(1,i-1)+1}$$

or

$$\langle 1_i, 0_k| \, a_k a_i^\dagger \, |0_i, 1_k\rangle = -(-1)^{\sum(i+1,k-1)}. \tag{65.7}$$

Comparing (65.7) with (65.5), we see that these quantities have opposite signs, i.e.

$$\widehat{a}_i\widehat{a}_k^\dagger + \widehat{a}_k\widehat{a}_i^\dagger = 0, \quad i \neq k.$$

For the diagonal matrix $\widehat{a}_i\widehat{a}_i^\dagger$, we find

$$\widehat{a}_i\widehat{a}_i^\dagger = 1 - N_i \tag{65.8}$$

Adding this to (65.6), we obtain

$$\widehat{a}_i\widehat{a}_i^\dagger + \widehat{a}_i^\dagger\widehat{a}_i = 1.$$

Both the above equations can be written in the form

$$\widehat{a}_i\widehat{a}_k^\dagger + \widehat{a}_k^\dagger\widehat{a}_i = \delta_{ik}. \tag{65.9}$$

On carrying out similar calculations, we find for the products $\widehat{a}_i\widehat{a}_k$ the relations

$$\widehat{a}_i\widehat{a}_k + \widehat{a}_k\widehat{a}_i = 0, \tag{65.10}$$

and in particular $\widehat{a}_i\widehat{a}_i = 0$.

Thus we see that the operators $\widehat{a}_i$ and $\widehat{a}_k$ (or $\widehat{a}_k^\dagger$) for $i \neq k$ anticommute, whereas in the case of Bose statistics they commuted with one another. This difference is perfectly natural. In the case of Bose statistics, the operators $\widehat{a}_i$ and $\widehat{a}_k$ were completely independent; each of the operators $\widehat{a}_i$ acted only on a single variable N_i, and the result of this action did not depend on the values of the other occupation numbers. In the case of Fermi statistics, however, the result of the action of the operator $\widehat{a}_i$ depends not only on the number N_i itself, but also on the occupation numbers of all the preceding states, as we see from the definition (65.4). Hence the action of the various operators $\widehat{a}_i, \widehat{a}_k$ cannot be considered independent.

The properties of the operators $\widehat{a}_i, \widehat{a}_i^\dagger$ having been thus defined, all the remaining formulae (64.13)–(64.18) remain valid. The formulae (64.23)–(64.25), which express the operators of physical quantities in terms of the ψ-operators defined by (64.20), also hold good. The commutation rules (64.21), (64.22), however, are now replaced by

$$\widehat{\psi}^\dagger(\xi')\widehat{\psi}(\xi) + \widehat{\psi}(\xi)\widehat{\psi}^\dagger(\xi') = \delta(\xi - \xi'), \tag{65.11}$$

$$\widehat{\psi}(\xi')\widehat{\psi}(\xi) + \widehat{\psi}(\xi)\widehat{\psi}(\xi') = 0. \tag{65.12}$$

If the system consists of particles of different kinds, second quantization operators must be defined for each kind of particle (as already mentioned at the end of §64). Operators belonging to bosons and fermions commute; those belonging to different fermions may formally be regarded as either commutative or anticommutative within the limits of non-relativistic theory. On either assumption the results obtained by means of the second quantization method are the same.

However, with a view to later applications in the relativistic theory, which allows different particles to be transformed into one another, we should assume that the creation and annihilation operators for different fermions anticommute. This becomes evident if we regard as "different" particles two different internal states of a single complex particle.

MATHEMATICAL APPENDICES

§ a. Hermite polynomials

The equation

$$y'' - 2xy' + 2ny = 0 \qquad\qquad (a.1)$$

belongs to a class which can be solved by what is called *Laplace's method.*[21])

This method is applicable to any linear equation of the form

$$\sum_{m=0}^{n}(a_m + b_m x)\frac{\mathrm{d}^m y}{\mathrm{d}x^m} = 0,$$

whose coefficients are of degree in x not higher than the first, and consists in the following procedure. We form the polynomials

$$P(t) = \sum_{m=0}^{n} a_m t^m, \quad Q(t) = \sum_{m=0}^{n} b_m t^m$$

and from them the function

$$Z(t) = \frac{1}{Q}\exp\int \frac{P}{Q}\mathrm{d}t,$$

which is determined to within a constant factor. Then the solution of the equation under consideration can be expressed as a complex integral:

$$y = \int_C Z(t)\mathrm{e}^{xt}\mathrm{d}t,$$

where the path of integration C is taken so that the integral is finite and non-zero, and the function

$$V = \mathrm{e}^{xt}QZ$$

returns to its original value when t describes the contour C (which may be either closed or open).

---

[21]) See, for instance, E. Goursat, *Cours d'Analyse Mathematique*, Vol. II, Gauthier-Villars, Paris; V. I. Smirnov, *Course of Higher Mathematics*, Vol. III, Part 2, Pergamon, Oxford, 1964.

In the case of equation (a.1) we have

$$P = t^2 + 2n; \quad Q = -2t, \quad Z = \frac{1}{2t^{n+1}} e^{-t^2/4}, \quad V = \frac{1}{t^n} e^{xt-t^2/4},$$

so that its solution is

$$y = \int \exp\left(xt - \frac{t^2}{4}\right) \frac{dt}{t^{n+1}}. \tag{a.2}$$

For physical applications we need only consider values $n > -1/2$. For these values the contour of integration can be taken as C_1 or C_2 (Fig. 12); these satisfy the required conditions[22], since the function V vanishes at their ends ($t = +\infty$ or $t = -\infty$).

Let us find the values of the parameter n for which equation (a.1) has solutions finite for all finite x, which tend to infinity, as $x \to \pm\infty$, not more rapidly than every finite power of x. First, we consider non-integral values of n. The integrals (a.2) along C_1 and C_2 then give two independent solutions of equation (a.1). We transform the integral along C_1 by introducing the variable u such that $t = 2(x - u)$. Omitting a constant factor, we find

$$y = e^{x^2} \int_{C_1'} \frac{e^{-u^2}}{(u-x)^{n+1}} du, \tag{a.3}$$

where the integration is taken over the contour C_1' in the complex plane of u, as shown in Fig. 13.
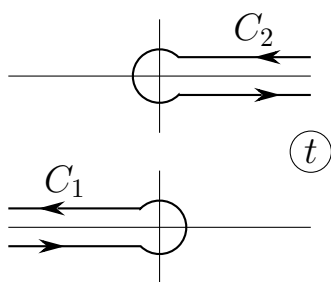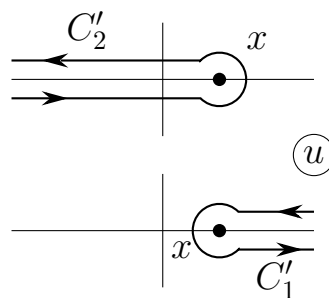

FIG. 52

FIG. 53

As $x \to +\infty$, the whole path of integration C_1' moves to infinity, and the integral in (a.3) tends to zero as e^{-x^2}. As $x \to -\infty$, however, the path

---

[22] These paths will not serve for negative integral n, since the integral (a.2) along them then vanishes identically.

of integration extends along the whole of the real axis, and the integral in (a.3) does not tend exponentially to zero, so that the function $y(x)$ becomes infinite essentially as e^{x^2} Similarly, it is easy to see that the integral (a.2) along the contour C_2 diverges exponentially as $x \to +\infty$.

For positive integral n (including zero), on the other hand, the integrals along the straight parts of the path of integration cancel, and the two integrals (a.3), along C_1' and C_2', reduce to an integral along a closed path round the point $u = x$. Thus we have the solution

$$y(x) = \mathrm{e}^{x^2} \oint \frac{\mathrm{e}^{-u^2}}{(u-x)^{n+1}} \mathrm{d}u,$$

which satisfies the conditions stated. According to Cauchy's well-known formula for the derivatives of an analytic function,

$$f^{(n)}(x) = \frac{n!}{2\pi \mathrm{i}} \oint \frac{f(t)}{(t-x)^{n+1}} \mathrm{d}t,$$

$y(x)$ is, apart from a constant factor, an *Hermite polynomial*:

$$H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2}. \tag{a.4}$$

The polynomial H_n, expanded in decreasing powers of x, has the open form

$$H_n(x) = (2x)^n - \frac{n(n-1)}{1}(2x)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{1 \cdot 2}(2x)^{n-4} - \ldots \tag{a.5}$$

It contains only powers of x which are of the same parity as n. We may write out here the first few Hermite polynomials:

$$H_0 = 1, H_1 = 2x, H_2 = 4x^2 - 2, H_3 = 8x^3 - 12x, H_4 = 16x^4 - 48x^2 + 12. \tag{a.6}$$

To calculate the normalization integral, we replace $\mathrm{e}^{-x} H_n$ by its expression in (a.4) and integrate n times by parts:

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2} H_n^2(x) \mathrm{d}x = \int_{-\infty}^{+\infty} (-1)^n H_n(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \mathrm{d}x = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \frac{\mathrm{d}^n H_n}{\mathrm{d}x^n} \mathrm{d}x.$$

But $\mathrm{d}^n H_n / \mathrm{d}x^n$ is a constant, $2^n n!$. Thus

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2} H_n^2(x) \mathrm{d}x = 2^n n! \sqrt{\pi}. \tag{a.7}$$

§ b. The Airy function

The equation

$$y'' - xy = 0 \tag{b.1}$$

is of Laplace's type (see §a). Following the general method, we form the functions

$$P = t^2, \quad Q = -1, \quad Z = -\exp(-t^3/3), \quad V = \exp(xt - t^3/3),$$

so that the solution can be represented in the form

$$y(x) = \text{const} \cdot \int_C \exp(xt - t^3/3)\mathrm{d}t, \tag{b.2}$$

The path of integration C must be chosen so that the function V vanishes at both ends of it. These ends must therefore go to infinity in the regions of the complex plane of t in which $\mathrm{Re}(t^3) > 0$ (the shaded regions in Fig. 14).
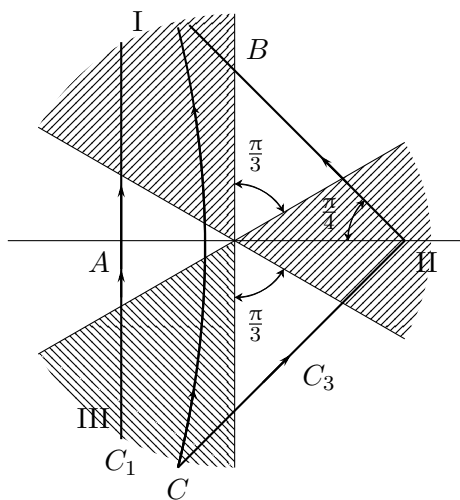
A solution finite for all x is obtained by taking the path C as shown in the figure. It can be displaced in any manner provided that the ends of it go to infinity in the same two shaded sectors (I and III in Fig. 14). We notice that, by taking a path which lay in sectors III and II (say), we should obtain a solution which becomes infinite as $x \to \infty$.

Deforming the path C so that it goes along the imaginary axis, we obtain the function (b.2) in the form (substituting $t = \mathrm{i}u$)

FIG.54

$$\Phi(x) = \frac{1}{\sqrt{x}} \int_0^\infty \cos\left(ux + \frac{u^3}{3}\right)\mathrm{d}u. \tag{b.3}$$

The constant in (b.2) has been put equal to $-\mathrm{i}/2\sqrt{\pi}$, and we have denoted the function thus obtained by $\Phi(x)$; it is called the *Airy function.*[23]

---

[23] We follow the definition proposed by V. A. Fok; see G. D. Yakovleva, *Tablitsy funktsiĭ Eriĭ* (*Tables of Airy Functions*), Nauka. Moscow, 1969. The function $\Phi(x)$ is one of two defined by Fok, who denotes it by $V(x)$. In the literature, another definition of the Airy function is also found, which differs from (b.3) by a constant factor: $\mathrm{Ai}(x) = \Phi(x)/\sqrt{\pi}$.

The asymptotic expression for $\Phi(x)$ for large values of x is obtained by calculating the integral (b.2) by the saddle-point method. For $x > 0$, the exponent in the integrand has an extremum for $t = \pm\sqrt{x}$, and the "direction of steepest descent" of the integrand is parallel to the imaginary axis. Accordingly, to obtain the asymptotic expression for large positive x, we expand the exponent in powers of $t + \sqrt{x}$ and integrate along the line C_1 (Fig. 14, which is parallel to the imaginary axis; the distance $OA = \sqrt{x}$. Making the substitution $t = -\sqrt{x} + iu$, we have

$$\Phi(x) \approx \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{2}{3}x^{3/2} - \sqrt{x}u^2\right) du,$$

whence

$$\Phi(x) \approx \frac{1}{2x^{1/4}} \exp\left(-\frac{2}{3}x^{3/2}\right). \tag{b.4}$$

Thus, for large positive x, the function $\Phi(x)$ diminishes exponentially.

To obtain the asymptotic expression for large negative values of x, we notice that, for $x < 0$, the exponent has an extremum for $t = i\sqrt{|x|}$ and $t = -i\sqrt{|x|}$, and the direction of steepest descent at these points is along lines at angles $-\pi/4$ and $\pi/4$ respectively to the real axis. Taking as the path of integration the broken line C_3 (the distance $OB = \sqrt{|x|}$), we have, after some simple transformations,

$$\Phi(x) = \frac{1}{|x|^{1/4}} \sin\left(\frac{2}{3}|x|^{3/2} + \frac{\pi}{4}\right). \tag{b.5}$$

Thus, in the region of large negative x, the function $\Phi(x)$ is oscillatory. We may mention that the first (and highest) maximum of the function $\Phi(x)$ is $\Phi(1.02) = 0.95$.

The Airy function can be expressed in terms of Bessel functions of order $1/3$. The equation (b.1), as can easily be seen, has the solution

$$\sqrt{x}Z_{1/3}\left(\frac{2}{3}x^{3/2}\right),$$

where $Z_{1/3}(x)$ is any solution of Bessel's equation of order $1/3$. The solution which is the same as (b.3) is

$$\Phi(x) = \frac{\sqrt{\pi x}}{3}\left[I_{-1/3}\left(\frac{2}{3}x^{3/2} - I_{1/3}\left(\frac{2}{3}x^{3/2}\right)\right)\right] \equiv \sqrt{\frac{x}{3\pi}}K_{1/3}\left(\frac{2}{3}x^{3/2}\right) \text{ for } x > 0,$$

$$\Phi(x) = \frac{\sqrt{\pi|x|}}{3}\left[J_{-1/3}\left(\frac{2}{3}|x|^{3/2}\right) + J_{1/3}\left(\frac{2}{3}|x|^{3/2}\right)\right] \text{ for } x < 0,$$
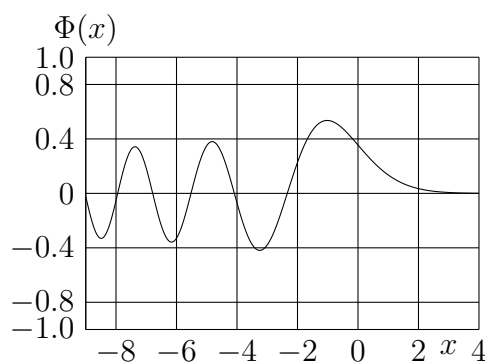
$$\tag{b.6}$$

where
$$I_\nu(x) = \mathrm{i}^{-\nu} J_\nu(\mathrm{i}x), \quad K_\nu(x) = \frac{\pi}{2\sin\nu\pi}\left[I_{-\nu}(x) - I_\nu(x)\right].$$

Using the recurrence relations

$$K_{\nu-1}(x) - K_{\nu+1}(x) = -\frac{2\nu}{x}K_\nu(x),$$

$$2K'_\nu(x) = -K_{\nu-1}(x) - K_{\nu+1}(x)$$

we   easily   find   for   the   derivative   of   the   Airy   function

$$\Phi'(x) = -\frac{x}{\sqrt{3\pi}}K_{2/3}\left(\frac{2}{3}x^{3/2}\right)$$

$$\text{for } x > 0. \quad \text{(b.7)}$$

When $x = 0$,

$$\Phi(0) = \frac{\sqrt{\pi}}{3^{2/3}\Gamma(2/3)} = 0.629,$$

$$\Phi'(0) = \frac{3^{1/6}\Gamma(2/3)}{2\sqrt{\pi}} = -0.459. \tag{b.8}$$

FIG.55
Figure. 15 shows a graph of the Airy function.

§ c. Legendre polynomials[24])

The *Legendre polynomials* $P_l(\cos\theta)$ are defined by the formula

$$P_l(\cos\theta) = \frac{1}{2^l l!}\frac{d^l}{(d\cos\theta)^l}(\cos^2\theta - 1)^l. \tag{c.1}$$

They satisfy the differential equation

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}P_l}{\mathrm{d}\theta}\right) + l(l+1)P_l = 0. \tag{c.2}$$

The associated Legendre polynomials are defined by

$$P_l^m(\cos\theta) = \sin^m\theta\frac{\mathrm{d}^m P_l(\cos\theta)}{(\mathrm{d}\cos\theta)^m} = \frac{1}{2^l l!}\sin^m\theta\frac{\mathrm{d}^{l+m}}{(\mathrm{d}\cos\theta)^{l+m}}(\cos^2\theta - 1)^l \tag{c.3}$$

---

[24]) There are in the mathematical literature many good accounts of the theory of spherical harmonics. Here we shall give, for reference, only a few basic relations, and make no attempt at a systematic discussion of the theory of these functions.

or, equivalently,

$$P_l^m(\cos\theta) = (-1)^m \frac{(l+m)!}{(l-m)!2^l l!} \sin^{-m}\theta \frac{\mathrm{d}^{l-m}}{(\mathrm{d}\cos\theta)^{l-m}}(\cos^2\theta - 1)^l, \qquad \text{(c.4)}$$

with $m = 0, 1, \ldots, l$. The associated polynomials satisfy the equation

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}P_l^m}{\mathrm{d}\theta}\right) + \left[l(l+1) - \frac{m^2}{\sin^2\theta}\right]P_l^m = 0. \qquad \text{(c.5)}$$

The normalization integral $\int_{-1}^{1}[P_l(\mu)]^2\,\mathrm{d}\mu(\mu = \cos\theta)$ for the Legendre polynomials is calculated by substituting (c.1) and integrating l times by parts, which gives

$$\frac{(-1)^l}{2^{2l}(l!)^2}\int_{-1}^{1}(\mu^2 - 1)^l\frac{\mathrm{d}^{2l}}{\mathrm{d}\mu^{2l}}(\mu^2 - 1)^l\mathrm{d}\mu = \frac{(2l)!}{2^{2l}(l!)^2}\int_{-1}^{1}(1 - \mu^2)^l\mathrm{d}\mu.$$

Substitution of $u = (1\mu)/2$ reduces this integral to Euler's beta function, and the result is

$$\int_{-1}^{1}[P_l(\mu)]^2\mathrm{d}\mu = \frac{2}{2l+1}. \qquad \text{(c.6)}$$

Similarly, it is easily seen that the functions $P_l(\mu)$ with different l are orthogonal:

$$\int_{-1}^{1}P_l(\mu)P_{l'}(\mu)\mathrm{d}\mu = 0, l \neq l'. \qquad \text{(c.7)}$$

The calculation of the normalization integral for the associated Legendre polynomials is easily effected by a similar method. We write $[P_l^m(\mu)]^2$ as a product of the expressions (c.3) and (c.4), and integrate lm times by parts; the result is

$$\int_{-1}^{1}[P_l^m(\mu)]^2\mathrm{d}\mu = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}. \qquad \text{(c.8)}$$

It is also easily seen that the functions P_l^m with different l (and the same m) are orthogonal:

$$\int_{-1}^{1}P_l^m(\mu)P_{l'}^m(\mu)\mathrm{d}\mu = 0, l \neq l'. \qquad \text{(c.9)}$$

The calculation of the integrals of products of three Legendre polynomials is discussed in §107.

The following *addition theorem* holds for Legendre polynomials. Let γ be the angle between two directions defined by the spherical angles θ, φ and

$\theta', \varphi' : \cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\varphi - \varphi')$. Then

$$P_l(\cos\gamma) = P_l(\cos\theta)P_l(\cos\theta') +$$

$$+ \sum_{m-1}^{l} 2\frac{(l-m)!}{(l+m)!}P_l^m(\cos\theta)P_l^m(\cos\theta')\cos[m(\varphi - \varphi')]. \quad \text{(c.10)}$$

This theorem can also be written in terms of the spherical harmonic functions defined by (28.7):

$$P_l(\boldsymbol{nn'}) = \frac{4\pi}{2l+1}\sum_{m=-1}^{l} Y_{lm}^*(\boldsymbol{n'})Y_{lm}(\boldsymbol{n}). \quad \text{(c.11)}$$

Here $\boldsymbol{n}$ and $\boldsymbol{n'}$ are two unit vectors, and $Y_{lm}(\boldsymbol{n})$ denotes the spherical harmonic function of the polar angle and azimuth of the direction of $\boldsymbol{n}$ relative to a fixed system of coordinates.

If equation (c.10) is multiplied by $P_{l'}(\cos\theta)$ and integrated over $\mathrm{d}o = \sin\theta\mathrm{d}\theta\mathrm{d}\varphi$, the integration with respect to φ gives zero for all terms on the right that contain factors $\cos[m(\varphi - \varphi')]$; using (c.6) and (c.7), we obtain

$$\int P_l(\cos\gamma)P_{l'}(\cos\theta)\mathrm{d}o = \delta_{ll'}\frac{4\pi}{2l+1}P_l(\cos\theta').$$

This result may be written in the symmetrical form

$$\int P_l(\boldsymbol{n_1n_2})P_{l'}(\boldsymbol{n_1n_3})\mathrm{d}o_1 = \delta_{ll'}\frac{4\pi}{2l+1}P_l(\boldsymbol{n_2n_3}), \quad \text{(c.12)}$$

where $\boldsymbol{n_1}, \boldsymbol{n_2}, \boldsymbol{n_3}$ are three unit vectors and the integration is with respect to the direction of $\boldsymbol{n_1}$.

Finally, we give the first few normalized spherical harmonics Y_{lm}:

$$Y_{00} = \frac{1}{\sqrt{4\pi}}, \quad Y_{10} = \mathrm{i}\sqrt{\frac{3}{4\pi}}\cos\theta, \quad Y_{1,\pm 1} = \pm\mathrm{i}\sqrt{\frac{3}{8\pi}}\sin\theta\cdot\mathrm{e}^{\pm\mathrm{i}\varphi},$$

$$Y_{20} = \sqrt{\frac{5}{16\pi}}(1 - 3\cos^2\theta), \quad Y_{2,\pm 1} = \pm\sqrt{\frac{15}{8\pi}}\cos\theta\sin\theta\cdot\mathrm{e}^{+\mathrm{i}\varphi},$$

$$Y_{2,\pm 2} = -\sqrt{\frac{15}{32\pi}}\sin^2\theta\cdot\mathrm{e}^{\pm 2\mathrm{i}\varphi}, \quad Y_{30} = -\mathrm{i}\sqrt{\frac{7}{16\pi}}\cos\theta(5\cos^2\theta - 3),$$

$$Y_{3,\pm 1} = \pm\mathrm{i}\sqrt{\frac{21}{64\pi}}\sin\theta(5\cos^2\theta - 1)\mathrm{e}^{\pm\mathrm{i}\varphi},$$

$$Y_{3,\pm 2} = -\mathrm{i}\sqrt{\frac{105}{32\pi}}\cos\theta\sin^2\theta\cdot\mathrm{e}^{\pm 2\mathrm{i}\varphi}, \quad Y_{3,\pm 3} = \pm\mathrm{i}\sqrt{\frac{35}{64\pi}}\sin^3\theta\cdot\mathrm{e}^{\pm 3\mathrm{i}\varphi}.$$

§ d. The confluent hypergeometric function

The *confluent hypergeometric function* is defined by the series

$$F(\alpha, \gamma, z) = 1 + \frac{\alpha}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{z^2}{2!} + \dots, \tag{d.1}$$

which converges for all finite z; the parameter α is arbitrary, while the parameter γ is supposed not zero or a negative integer. If α is a negative integer (or zero), $F(\alpha, \gamma, z)$ reduces to a polynomial of degree $|\alpha|$.

The function $F(\alpha, \gamma, z)$ satisfies the differential equation

$$zu'' + (\gamma - z)u' - \alpha u = 0, \tag{d.2}$$

as is easily seen by direct verification.[25] By the substitution $u = z^{1-\gamma}u_1$, this equation is transformed into another of the same form,

$$zu_1'' + (2 - \gamma - z)u_1' - (\alpha - \gamma + 1)u_1 = 0. \tag{d.3}$$

Hence we see that, for non-integral γ, equation (d.2) has also the particular integral $z^{1-\gamma}F(\alpha - \gamma + 1, 2 - \gamma, z)$, which is linearly independent of (d.1), so that the general solution of equation (d.2) is of the form

$$u = c_1 F(\alpha, \gamma, z) + c_2 z^{1-\gamma} F(\alpha - \gamma + 1, 2 - \gamma, z). \tag{d.4}$$

The second term, unlike the first, has a singular point at $z = 0$.

Equation (d.2) is of Laplace's type, and its solutions can be represented as contour integrals. Following the general method, we form the functions

$$P(t) = \gamma t - \alpha, \quad Q(t) = t(t-1), \quad Z(t) = t^{\alpha-1}(t-1)^{\gamma-\alpha-1},$$

so that

$$u = \int e^{tz} t^{\alpha-1}(t-1)^{\gamma-\alpha-1} dt. \tag{d.5}$$

The path of integration must be chosen so that the function $V(t) = e^{tz} t^{\alpha}(t-1)^{\gamma-\alpha}$ returns to its original value on traversing the path. Applying the same method to equation (d.3), we can obtain for u a contour integral of another form:

$$u = z^{1-\gamma} \int e^{tz} t^{\alpha-\gamma}(t-1)^{-\alpha} dt.$$

---

[25] The equation (d.2) with a negative integral γ does not require special discussion, since it can be reduced to a case of positive integral γ by the transformation which gives equation (d.3).

The substitution $tz \to t$ reduces this integral to the convenient form

$$u(z) = \int \mathrm{e}^t (t-z)^{-\alpha} t^{\alpha-\gamma} \mathrm{d}t, \tag{d.6}$$

and the corresponding function V to

$$V(t) = \mathrm{e}^t (t-z)^{-\alpha} t^{\alpha-\gamma} \mathrm{d}t,$$

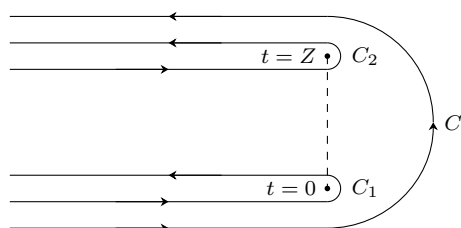The integrand in (d.6) has in general two singular points, at $t = z$ and $t = 0$. We take a contour of integration C which passes from infinity ($\mathrm{Re}\, t \to -\infty$) round the two singular points in the positive direction and back to infinity (Fig. 16). This contour satisfies the required conditions, since $V(t)$ vanishes at its ends. The integral (d.6), taken along the contour C, has no singular point for $z = 0$; hence it must be the same, apart from a constant factor, as the function $F(\alpha, \gamma, z)$, which also has no singularity. For $z = 0$ the two singular points of the integrand coincide; according to a well-known formula in the theory of the gamma function,

FIG. 56

$$\frac{1}{2\pi \mathrm{i}} \int_C \mathrm{e}^t t^{-\gamma} \mathrm{d}t = \frac{1}{\Gamma(\gamma)}. \tag{d.7}$$

Since $F(\alpha, \gamma, 0) = 1$, it is evident that

$$F(\alpha, \gamma, z) = \frac{\Gamma(\gamma)}{2\pi \mathrm{i}} \int_C \mathrm{e}^t (t-z)^{-\alpha} t^{\alpha-\gamma} \mathrm{d}t. \tag{d.8}$$

The integrand in (d.5) has singular points at $t = 0$ and $t = 1$. If $\mathrm{Re}(\gamma - \alpha) > 0$, and γ is not a positive integer, the path of integration can be taken as a contour C' starting from the point $t = 1$, passing round the point $t = 0$ in the positive direction, and returning to $t = 1$ (Fig. 17); for $\mathrm{Re}(\gamma - \alpha) > 0$, the function $V(t)$ returns to its original value of zero on passing round such a contour.[26] The integral thus defined again has no singularity for $z = 0$, and is related to $F(\alpha, \gamma, z)$ by

$$F(\alpha, \gamma, z) = -\frac{1}{2\pi \mathrm{i}} \frac{\Gamma(1-\alpha)\Gamma(\gamma)}{\Gamma(\gamma-\alpha)} \oint_{C'} \mathrm{e}^{tz} (-t)^{\alpha-1} (1-t)^{\gamma-\alpha-1} \mathrm{d}t. \tag{d.9}$$

---

[26] If γ is a positive integer, C' can be any contour which passes round both the points $t = 0$ and $t = 1$

The following remark should be made concerning the integrals (d.8), (d.9). For non-integral α and γ, the integrands are not one-valued functions. Their values at each point are supposed chosen in accordance with the condition that the complex quantity which is raised to a power is taken with the argument whose absolute value is least.
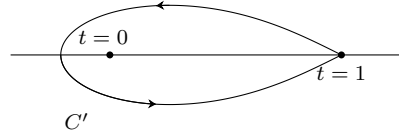

FIG. 57

We may notice the useful relation

$$F(\alpha, \gamma, z) = e^z F(\gamma - \alpha, \gamma, -z), \tag{d.10}$$

which is obtained at once by substituting $t \to t + z$ in the integral (d.8).

We have already remarked that, if $\alpha = -n$, where n is a positive integer, the function $F(\alpha, \gamma, z)$ reduces to a polynomial. A concise formula can be obtained for these polynomials. Making in the integral (d.9) the substitution $t \to 1 - (t/z)$ and applying Cauchy's formula to the resulting integral, we find

$$F(-n, \gamma, z) = \frac{1}{\gamma(\gamma + 1) \dots (\gamma + n - 1)} z^{1-\gamma} e^z \frac{d^m}{dz^n}(e^{-z} z^{\gamma + n - 1}). \tag{d.11}$$

If also $\gamma =$ a positive integer m, we have the formula

$$F(-n, m, z) = \frac{(-1)^{m-1}}{m(m+1) \dots (m+n-1)} e^z \frac{d^{m+n-1}}{dz^{m+n-1}}(e^{-z} z^n). \tag{d.12}$$

This formula is obtained by applying Cauchy's formula to the integral derived from (d.8) by the substitution $t \to z - t$.

The polynomials $F(-n, m, z), 0 \leqslant m \leqslant n$, are (apart from a constant factor) the *generalized Laguerre polynomials*

$$L_n^m(z) = (-1)^m \frac{(n!)^2}{m!(n-m)!} F(-(n-m), m+1, z) =$$

$$= \frac{n!}{(n-m)!} e^z \frac{d^n}{dz^n} e^{-z} z^{n-m} = (-1)^m \frac{n!}{(n-m)!} e^z z^{-m} \frac{d^{n-m}}{dz^{n-m}} e^{-z} z^n. \tag{d.13}$$

The polynomials L_n^m for $m = 0$ are denoted by $L_n(z)$ and are called simply *Laguerre polynomials*; from (d.13) we have

$$L_n)z = e^z \frac{d^n}{dz^n}(e^{-z} z^n).$$

The integral representation (d.8) is convenient for obtaining the asymptotic expansion of the confluent hypergeometric function for large z. We

deform the contour into two contours C_1 and C_2 (Fig. 56), which pass round the points $t = 0$ and $t = z$ respectively; the lower branch of C_2 and the upper branch of C_1 are supposed to join at infinity. To obtain an expansion in inverse powers of z, we take $(-z)^{-\alpha}$ outside the parenthesis in the integrand. In the integral along the contour C_2, we make the substitution $t \to t + z$; the contour C_2 is thereby transformed into C_1 We thus represent the formula (d.8) as

$$F(\alpha, \gamma, z) = \frac{\Gamma(\gamma)}{\Gamma(\gamma - \alpha)}(-z)^{-\alpha}G(\alpha, \alpha - \gamma + 1, -z)+$$
$$+ \frac{\Gamma(\gamma)}{\Gamma(\alpha)}e^z z^{\alpha - \gamma}G(\gamma - \alpha, 1 - \alpha, z), \quad \text{(d.14)}$$

where

$$G(\alpha, \beta, z) = \frac{\Gamma(1 - \beta)}{2\pi i}\int_{C_1}\left(1 + \frac{t}{z}\right)^{-\alpha}t^{\beta - 1}e^t dt. \qquad \text{(d.15)}$$

In raising $-z$ and z to powers in the formula (d.14) we must take the arguments which have the smallest absolute value. Finally, expanding $(1 + t/z)^{-\alpha}$ in the integrand in powers of t/z and applying formula (d.7), we have for $G(\alpha, \beta, z)$ the asymptotic series

$$G(\alpha, \beta, z) = 1 + \frac{\alpha\beta}{1!z} + \frac{\alpha(\alpha + 1)\beta(\beta + 1)}{2!z^2} + \dots \qquad \text{(d.16)}$$

Formulae (d.14) and (d.16) give the asymptotic expansion of the function $F(\alpha, \gamma, z)$.

   For positive integral γ, the second term in the general solution (d.4) of equation (d.2) is either the same as the first term (if $\gamma = 1$) or meaningless (if $\gamma > 1$). In this case we can take, as a set of two linearly independent solutions, the two terms in formula (d.14), i.e. the integrals (d.8) taken along the contours C_1 and C_2 (these contours, like C, satisfy the required conditions, so that the integrals along them are solutions of equation (d.2)). The asymptotic form of these solutions is given by the formulae already obtained; it remains for us to find their expansion in ascending powers of z. To do this, we start from equation (d.14) and the analogous equation for the function $z^{1-\gamma}F(\alpha - \gamma + 1, 2 - \gamma, z)$. From these two equations we express $G(\alpha, \alpha - \gamma + 1, -z)$ in terms of $F(\alpha, \gamma, z)$ and $F(\alpha - \gamma + 1, 2 - \gamma, z)$; we then put $\gamma = p + \varepsilon(p$ being a positive integer), and pass to the limit $\varepsilon \to 0$, resolving the indeterminacy by L'Hospital's rule. A fairly lengthy calculation

gives the following expansion:

$$G(\alpha, \alpha - p + 1, -z) = \frac{\sin \pi\alpha \cdot \Gamma(p - a)}{\pi\Gamma(p)} z^{\alpha} \times$$

$$\times \left\{ \ln z \cdot F(\alpha, p, z) + \sum_{s=0}^{\infty} \frac{\Gamma(p)\Gamma(\alpha + s)\left[\psi(\alpha + s) - \psi(p + s) - \psi(s + 1)\right]}{\Gamma(\alpha)\Gamma(s + p)\Gamma(s + 1)} z^{s} + \right.$$

$$\left. + \sum_{s=1}^{p-1} (-1)^{s+1} \frac{\Gamma(s)\Gamma(\alpha - s)\Gamma(p)}{\Gamma(\alpha)\Gamma(p - s)} z^{-s} \right\}, \quad \text{(d.17)}$$

where ψ denotes the logarithmic derivative of the gamma function: $\psi(\alpha = \Gamma'(\alpha)/\Gamma(\alpha)$.

§ e. The hypergeometric function

The *hypergeometric function* is defined in the circle $|z| < 1$ by the series

$$F(\alpha, \beta, \gamma, z) = 1 + \frac{\alpha\beta}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha + 1)\beta(\beta + 1)}{\gamma(\gamma + 1)}\frac{z^2}{2!} + \dots, \quad \text{(e.1)}$$

and for $|z| > 1$ it is obtained by analytical continuation of this series (see (e.6)). The hypergeometric function is a particular integral of the differential equation

$$z(1 - z)u'' + [\gamma - (\alpha + \beta + 1)z]u' - \alpha\beta u = 0. \quad \text{(e.2)}$$

The parameters α and β are arbitrary, while $\neq 0, -1, -2, \dots$ The function $F(\alpha, \beta, \gamma, z)$ is evidently symmetrical with respect to the parameters α and β.[27] The second independent solution of equation (e.2) is

$$z^{1-\gamma}F(\beta - \gamma + 1, \alpha - \gamma + 1, 2 - \gamma, z);$$

it has a singular point at $z = 0$.

We shall give here for reference a number of relations obeyed by the hypergeometric function.

---

[27] The confluent hypergeometric function is obtained from $F(\alpha, \beta, \gamma, z)$ by taking the limit

$$F(\alpha, \gamma, z) = \lim_{\beta \to \infty} F(\alpha, \beta, \gamma, \frac{z}{\beta})$$

The notation $_2F_1(\alpha, \beta, \gamma, z)$ for the hypergeometric function and $_1F_1(\alpha, \gamma, z)$ for the confluent hypergeometric function is also used in the literature. The subscripts to the left and right of F show the numbers of parameters in the numerators and denominators respectively of the terms in the series.

The function $F(\alpha, \beta, \gamma, z)$ can be represented for all z, if $\mathrm{Re}() > 0$, as an integral:

$$F(\alpha, \beta, \gamma, z) =$$
$$= -\frac{1}{2\pi\mathrm{i}} \frac{\Gamma(1-\alpha)\Gamma(\gamma)}{\Gamma(\gamma-\alpha)} \oint_{C'} (-t)^{\alpha-1}(1-t)^{\gamma-\alpha-1}(1-tz)^{-\beta}\mathrm{d}t, \quad \text{(e.3)}$$

taken along the contour C' shown in Fig. 17. That this integral in fact satisfies equation (e.2) is easily seen by direct substitution; the constant factor is chosen so as to give unity for $z = 0$.

The substitution $u = (1-z)^{\gamma-\alpha-\beta}u_1$ in equation (e.2) leads to an equation of the same form, with parameters $\gamma - \alpha, \gamma - \beta, \gamma$ in place of α, β, γ respectively. Hence we have

$$F(\alpha, \beta, \gamma, z) = (1-z)^{\gamma-\alpha-\beta}F(\gamma-\alpha, \gamma-\beta, \gamma, z) \qquad \text{(e.4)}$$

both sides of this equation satisfy the same equation, and they have the same value for $z = 0$.

The substitution $t \to t/(1 - z + zt)$ in the integral (e.3) leads to the following relation between hypergeometric functions with variables z and $z/(z1)$:

$$F(\alpha, \beta, \gamma, z) = (1-z)^{-\alpha}F\left(\alpha, \gamma-\beta, \gamma, \frac{z}{z-1}\right). \qquad \text{(e.5)}$$

The value of the many-valued expression $(1-z)^{-\alpha}$ in this formula (and of similar expressions in all the following formulae) is determined by the condition that the complex quantity which is raised to a power is taken with the argument whose absolute value is least.

Next we shall give, without proof, an important formula relating hypergeometric functions with variables z and $1/z$:

$$F(\alpha, \beta, \gamma, z) = \frac{\Gamma(\gamma)\Gamma(\beta-\alpha)}{\Gamma(\beta)\Gamma(\gamma-\alpha)}(-z)^{-\alpha}F\left(\alpha, \alpha+1-\gamma, \alpha+1-\beta, \frac{1}{z}\right) +$$
$$+ \frac{\Gamma(\gamma)\Gamma(\alpha-\beta)}{\Gamma(\alpha)\Gamma(\gamma-\beta)}(-z)^{-\beta}F\left(\beta, \beta+1-\gamma, \beta+1-\alpha, \frac{1}{z}\right). \quad \text{(e.6)}$$

This formula expresses $F(\alpha, \beta, \gamma, z)$ as a series which converges for $|z| > 1$, i.e. it is the analytical continuation of the original series (e.1).

The formula

$$F(\alpha, \beta, \gamma, z) = \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}F(\alpha, \beta, \alpha+\beta+1-\gamma, 1-z) +$$
$$+ \frac{\Gamma(\gamma)\Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha)\Gamma(\beta)}(1-z)^{\gamma-\alpha-\beta}F(\gamma-\alpha, \gamma-\beta, \gamma+1-\alpha-\beta, 1-z).$$
$$\text{(e.7)}$$

relates hypergeometric functions of z and $1-z$; again, the proof will not be given here. Combining (e.7) with (e.6), we obtain the relations

$$F(\alpha, \beta, \gamma, z) = \frac{\Gamma(\gamma)\Gamma(\beta-\alpha)}{\Gamma(\beta)\Gamma(\gamma-\alpha)}(1-z)^{-\alpha}F\left(\alpha, \gamma-\beta, \alpha+1-\beta, \frac{1}{1-z}\right) +$$

$$+ \frac{\Gamma(\gamma)\Gamma(\alpha-\beta)}{\Gamma(\alpha)\Gamma(\gamma-\beta)}(1-z)^{-\beta}F(\beta, \gamma-\alpha, \beta+1-\alpha, \frac{1}{1-z}), \quad \text{(e.8)}$$

$$F(\alpha, \beta, \gamma, z) = \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\beta)\Gamma(\gamma-\alpha)} \times$$

$$\times z^{-\alpha}F\left(\alpha, \alpha+1-\gamma, \alpha+\beta+1-\gamma, \frac{z-1}{z}\right) + \frac{\Gamma(\gamma)\Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha)\Gamma(\beta)} \times$$

$$\times (1-z)^{\gamma-\alpha-\beta}z^{\beta-\gamma}F\left(1-\beta, \gamma-\beta, \gamma+1-\alpha-\beta, \frac{z-1}{z}\right). \quad \text{(e.9)}$$

Each of the terms in the sums on the right of equations (e.6)–(e.9) is itself a solution of the hypergeometric equation.

If α (or β) is a negative integer or zero, $\alpha = -n$, the hypergeometric function reduces to a polynomial of the nth degree, and can be represented in the form

$$F(-n, \beta, \gamma, z) = \frac{z^{1-\gamma}(1-z)^{\gamma+n-\beta}}{\gamma(\gamma+1)\ldots(\gamma+n-1)}\frac{\mathrm{d}^n}{\mathrm{d}z^n}\left[z^{\gamma+n-1}(1-z)^{\beta-\gamma}\right]. \quad \text{(e.10)}$$

These polynomials are the same, apart from a constant factor, as the *Jacobi polynomials*, defined by

$$P_n^{(a,b)}(z) = \frac{(a+1)(a+2)\ldots(a+n)}{n!}F\left(-n, a+b+n+1, a+1, \frac{1-z}{2}\right) =$$

$$= \frac{(-1)^n}{2^n n!}(1-z)^{-a}(1+z)^{-b}\frac{\mathrm{d}^n}{\mathrm{d}z^n}\left[(1-z)^{a+n}(1+z)^{b+n}\right]. \quad \text{(e.11)}$$

For $a = b = 0$, the Jacobi polynomials are the Legendre polynomials. For $n = 0$, $P_0^{(a,b)} = 1$.

§ f. The calculation of integrals containing confluent hypergeometric functions

Let us consider an integral of the form

$$J_{\alpha\gamma}^{\nu} = \int_0^{\infty} \mathrm{e}^{-\lambda z}z^{\nu}F(\alpha, \gamma, kz)\mathrm{d}z. \quad \text{(f.1)}$$

We assume that it converges. If this is so we must have $\mathrm{Re}\,\nu > -1$ and $\mathrm{Re}\,\lambda > |\mathrm{Re}\,k|$; if α is a negative integer, the latter condition can be replaced by $\mathrm{Re}\,\lambda > 0$. Using for $F(\alpha, \gamma, kz)$ the integral representation (d.9) and effecting the integration over z under the contour integral, we have

$$J_{\alpha\gamma}^{\nu} = -\frac{1}{2\pi i}\frac{\Gamma(1-\alpha)\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}\lambda^{-\nu-1}\Gamma(\nu+1)\times$$
$$\times \oint_{C'} (-t)^{\alpha-1}(1-t)^{\gamma-\alpha-1}(1-(k/\lambda)t)^{-\nu-1}\mathrm{d}t.$$

Using (e.3), we have finally

$$J_{\alpha\gamma}^{\nu} = \Gamma(\nu+1)\lambda^{-\nu-1}F(\alpha, \nu+1, \gamma, k/\lambda). \tag{f.2}$$

In the cases where the function $F(\alpha, \nu+1, \gamma, k/\lambda)$ reduces to a polynomial, we have for the integral $J_{\alpha\gamma}^{\nu}$ an expression in terms of elementary functions:

$$J_{\alpha\gamma}^{\gamma+n-1} = (-1)^n\Gamma(\gamma)\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\left[\lambda^{\alpha-\gamma}(\lambda-k)^{-\alpha}\right], \tag{f.3}$$

$$J_{-n\gamma}^{\nu} = (-1)^n\frac{\Gamma(\nu+1)(\lambda+k)^{\gamma+n-\nu-1}}{\gamma(\gamma+1)\dots(\gamma+n-1)}\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\left[\lambda^{-\nu-1}(\lambda-k)^{\nu-\gamma+1}\right], \tag{f.4}$$

$$J_{\alpha m}^{n} = \frac{(-1)^{m-n}}{k^{m-1}(1-\alpha)(2-\alpha)\dots(m-1-\alpha)}\times$$
$$\times\left\{-(m-1)!\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\left[\lambda^{\alpha-1}(\lambda-k)^{m-\alpha-1}\right] + \right.$$
$$+ n!(m-n-1)\dots(m-1)\lambda^{\alpha-n-1}(\lambda-k)^{-1+m-n-\alpha}\times$$
$$\left.\times\frac{\mathrm{d}^{m-n-2}}{\mathrm{d}\lambda^{m-n-2}}\left[\lambda^{m-\alpha-1}(\lambda-k)^{\alpha-1}\right]\right\}; \tag{f.5}$$

here m, n are integers, with $0 \leqslant n \leqslant m-2$.

Next, let us calculate the integral

$$J_{\nu} = \int_0^{\infty} \mathrm{e}^{-kz}z^{\nu-1}\left[F(-n, \gamma, kz)\right]^2\mathrm{d}z, \tag{f.6}$$

where n is an integer and $\mathrm{Re}\,v > 0$. To calculate this, we begin with a more general integral having $\mathrm{e}^{-\lambda z}$ instead of e^{-ks} in the integrand. We write one of the functions $F(-n, \gamma, kz)$ as a contour integral (d.9), and then integrate

over z, using formula (f.3):

$$\int_0^\infty = \mathrm{e}^{-\lambda z} z^{\nu-1} \left[F(-n,\gamma,kz)\right]^2 \mathrm{d}z = -\frac{1}{2\pi\mathrm{i}}(-1)^n \frac{\Gamma(1+n)\Gamma^2(\gamma)\Gamma(\nu)}{\Gamma^2(\gamma+n)} \times$$

$$\times \oint_{C'} (\lambda - kt - k)^{\gamma+n-\nu}(-t)^{-n-1}(1-t)^{\gamma+n-1} \times$$

$$\times \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\left[(\lambda-kt)^{-\nu}(\lambda-kt-k)^{\nu-\gamma}\right]\mathrm{d}t.$$

The nth derivative with respect to λ can evidently be replaced by a derivative of the same order with respect to t; we then put $\lambda = k$, and thereby return to the integral J_ν:

$$J_\nu = -\frac{1}{2\pi\mathrm{i}} \frac{\Gamma(n+1)\Gamma(\nu)\Gamma^2(\gamma)}{\Gamma^2(\gamma+n)k^\nu} \times$$

$$\times \oint_{C'} (-t)^{\gamma-\nu-1}(1-t)^{\gamma+n-1} \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left[(1-t)^{-\nu}(-t)^{\nu-\gamma}\right]\mathrm{d}t.$$

By integrating n times by parts, we transfer the operator $\mathrm{d}^n/\mathrm{d}t^n$ to the expression $(-t)^{\gamma-\nu-1}(1-t)^{\gamma+n-1}$, and then expand the derivative by Leibniz' formula. As a result, we obtain a sum of integrals, each of which reduces to Euler's well-known integral. We finally have the following expression for the integral required:

$$J_\nu = \frac{\Gamma(\nu)n!}{k^\nu\gamma(\gamma+1)\dots(\gamma+n-1)} \times$$

$$\times \left\{1 + \sum_{s=0}^{n-1} \frac{n(n-1)\dots(n-s)(\gamma-\nu-s-1)(\gamma-\nu-s)\dots(\gamma-\nu+s)}{[(s+1)!]^2\gamma(\gamma+1)\dots(\gamma+s)}\right\}.$$
$$\tag{f.7}$$

It is easy to see that the integrals J_ν are related by

$$J_{\gamma+p} = \frac{(\gamma-p-1)(\gamma-p)\dots(\gamma+p-1)}{k^{2p+1}} J_{\gamma-1-p}, \tag{f.8}$$

where p is any integer.

We similarly calculate the integral

$$J = \int_0^\infty \mathrm{e}^{-\lambda z} z^{\gamma-1} F(\alpha,\gamma,kz) F(\alpha',\gamma,k'z)\mathrm{d}z. \tag{f.9}$$

We represent the function $F(\alpha', \gamma, k'z)$ as a contour integral (d.9), and integrate over z, using formula (f.3) with $n = 0$:

$$J = -\frac{1}{2\pi i} \frac{\Gamma(1-\alpha')\Gamma^2(\gamma)}{\Gamma(\gamma-\alpha')} \oint_{C'} (-t)^{\alpha'-1}(1-t)^{\gamma-\alpha'-1}(\lambda - k't)^{\alpha-\gamma}\times$$
$$\times (\lambda - k't - k)^{-\alpha}dt.$$

By the substitution $t \to \lambda t/(k't + \lambda - k')$, this integral is brought to the form (e.3), giving

$$J = \Gamma(\gamma)\lambda^{\alpha+\alpha'-\gamma}(\lambda-k)^{-\alpha}(\lambda-k')^{-\alpha'} F\left(\alpha, \alpha', \gamma, \frac{kk'}{(\lambda-k)(\lambda-k')}\right). \quad \text{(f.10)}$$

If α (or α') is a negative integer, $\alpha = -n$, this expression can be rewritten, using (e.7), as

$$J = \frac{\Gamma^2(\gamma)\Gamma(\gamma+n-\alpha')}{\Gamma(\gamma+n)\Gamma(\gamma-\alpha')}\lambda^{-n+\alpha'-\gamma}(\lambda-k)^n(\lambda-k')^{-\alpha'}\times$$
$$\times F(-n, \alpha', -n+\alpha'+1-\gamma, \frac{\lambda(\lambda-k-k')}{(\lambda-k)(\lambda-k')}). \quad \text{(f.11)}$$

Finally, let us consider integrals of the form

$$J_\nu^{sp}(\alpha, \alpha') = \int_0^\infty \exp\left(-\frac{k+k'}{2}z\right) z^{\gamma-1+s} F(\alpha, \gamma, kz) F(\alpha', \gamma - p, k'z) dz. \tag{f.12}$$

The values of the parameters are supposed such that the integral converges absolutely; s and p are positive integers. The simplest of these integrals, $J_\gamma^{00}(\alpha, \alpha')$, is by (f.10),

$$J_\nu^{00}(\alpha, \alpha') = 2^\gamma \Gamma(\gamma)(k+k')^{\alpha+\alpha'-\gamma}(k'-k)^{-\alpha}\times$$
$$\times (k-k')^{-\alpha'} F\left(\alpha, \alpha', \gamma, -\frac{4kk'}{(k'-k)^2}\right), \quad \text{(f.13)}$$

if α (or α') is a negative integer, $\alpha = -n$, we can also write, by (f.11),

$$J_\gamma^{00}(-n, \alpha') = 2^\gamma \frac{\Gamma(\gamma)(\gamma-\alpha')(\gamma-\alpha'+1)\dots(\gamma-\alpha'+n-1)}{\gamma(\gamma+1)\dots(\gamma+n-1)}\times$$
$$\times (-1)^n (k+k')^{-n+\alpha'-\gamma}(k-k')^{n-\alpha'}\times$$
$$\times F\left[-n, \alpha', \alpha'+1-n-\gamma, \left(\frac{k+k'}{k-k'}\right)^2\right]. \quad \text{(f.14)}$$

The general formula for $J_\gamma^{sp}(\alpha, \alpha')$ can be derived, but it is so complex that it cannot be used conveniently. It is more convenient to use recurrence formulae, which enable us to reduce the integrals $J_\gamma^{sp}(\alpha, \alpha')$ to the integral with $s = p = 0$. The formula

$$J_\gamma^{sp}(\alpha, \alpha') = \frac{\gamma - 1}{k}\{J_{\gamma-1}^{s,p-1}(\alpha, \alpha') - J_{\gamma-1}^{s,p-1}(\alpha - 1, \alpha')\} \qquad \text{(f.15)}$$

enables us to reduce $J_\gamma^{sp}(\alpha, \alpha')$ to the integral with $p = 0$. The formula

$$J_\gamma^{s+1,0}(\alpha, \alpha') = \frac{4}{k^2 - k'^2}\{[\gamma(k - k')/2 - k\alpha + k'\alpha - k's]J_\gamma^{s0}(\alpha, \alpha')+$$
$$+ s(\gamma - 1 + s - 2\alpha')J_\gamma^{s-1,0}(\alpha, \alpha') + 2\alpha' s J_\gamma^{s-1,0}(\alpha, \alpha + 1)\} \quad \text{(f.16)}$$

then makes possible the final reduction to the integral with $s = p = 0$.[28]

---

[28] See W. Gordon, *Annalen der Physik* [5] 2, 1031, 1929.