AN
INTRODUCTION
TO
ELEMENTARY
PARTICLES

W. S. C. WILLIAMS

An

INTRODUCTION

to

ELEMENTARY PARTICLES

W. S. C. WILLIAMS

Clarendon Laboratory, University of Oxford
Oxford, England

1961

PREFACE

This book has been written for research workers in the field of experimental elementary particle physics and is intended to present an introduction to the theoretical methods and ideas which are used to describe the behaviour of elementary particles.

In the last fifteen years there has been a great increase in the use of symmetry properties, and of their associated conservation laws, in the field of elementary particles. These subjects are very important for they give considerable insight into the behaviour of elementary particles without requiring an extensive background of quantum mechanics for their understanding. At the time this book was started there was an obvious need for a book which covered these aspects of elementary particle physics in a manner which would appeal to experimental research workers.

The structure of the book follows my own preference as to the order in which a reader could become familiar with the subjects. In all cases the treatment is as elementary as is consistent with clarity and no claim is made to be rigorous. I felt that I had to draw a line beyond which the book would not go. It was chosen so as to avoid the calculation of transition matrix elements, although use is made of the transition rate formula as a means of comparing cross sections or lifetimes. The inclusion of the calculation of simple matrix elements would merely have covered the material of Fermi's excellent little book, "Elementary Particles". To have gone as far as the covariant calculation of matrix elements would have increased the size of the book to an unreasonable extent. However, field theory plays an important part in the theory of elementary particles and therefore a descriptive chapter on this subject is included. This allows us to use Feynman diagrams and to talk about propagators and vertices without involving the reader in involved matrix element calculations. Dispersion

v

theory is covered briefly but the book is too early to do more than mention the Mandelstam representation and its developments.

I wish to acknowledge the encouragement given to me by Professor J. C. Gunn at the critical stage in the writing of this book. I am indebted to Professor P. I. Dee for the opportunity to write the book and for the encouraging atmosphere in his department, where much of the writing was done. I am also indebted to Professor J. C. Gunn, Professor E. H. Bellamy, Drs. B. H. Bransden, K. Burton, G. M. Lewis, R. G. Moorhouse, D. T. Stewart for reading parts of the manuscript. The first half of the book was the material for a series of graduate lectures and I am grateful for helpful comments and criticisms received from the students. My thanks are due to Miss Ailie Currie for her excellent typing work and to my wife for invaluable help in checking the final typescript.

W. S. C. Williams

Department of Natural Philosophy
The University, Glasgow, Scotland.
February, 1961

Acknowledgments

CONTENTS

CHAPTER I

INTRODUCTION

I.I Preamble

The primary objective of this book is to serve as an introduction to the interpretation of some of the phenomena associated with elementary particles. We do not need to confuse the situation by debating the meaning of elementary in this respect: for us the term covers particles which are not, within present experience, bound states of other particles. Thus, the proton is an elementary particle, whereas the deuteron is not. To reduce the complication we only consider processes in which the number of particles in the initial state plus the number in the final state is four or less; this prevents the inclusion of complicated many-body processes and enhances the essential features. An example of this restriction is found in our treatment of the photoproduction of mesons. The description of the production from hydrogen contains the most important features for this book; production from complex nuclei is blurred by the presence of many nucleons and is not considered, except to mention production from deuterium. There is an excuse in this case, since the deuteron is the nearest we can approach to a free neutron target. We shall be discussing β decay which does involve complex nuclei; however, this case can fall within our requirement of four or less particles in the initial plus final state if we do not have to concern ourselves with the effects of nuclear structure.

Our treatment of processes is done in the most general way and for the greater part without reference to any specific models other than symmetry. By model we mean any theory or postulate as to the nature of the particles or of the forces involved. Such models come from quantum field theory or by suggesting a simple analytic form for the potential existing between particles. It is possible to obtain a

1

considerable amount of information about the particles not by appealing to such methods, but by the use of various symmetry laws. The description of the application of these laws is the main subject of this book. The material of Chapter IX is perhaps an exception to this restriction for it concerns field theory and the interaction of fields. However, the inclusion of this chapter subsequently allows us to discuss weak interactions (for which symmetries are important) in a more complete manner than would otherwise be possible.

Our plan is to develop the background necessary for the application of these symmetry laws; this will be done in the most elementary way possible and may be illustrated by direct example of actual physical processes. No rigour is claimed for any part of this book and the reader is warned that a more complete approach may be more subtle than is implied in this book; thus the theory of angular momentum can only reach its full development in group theory.

This first chapter is devoted to presenting some of the matter which is the basis of later chapters. We start by a brief statement of the postulates of quantum mechanics followed by a discussion of their implications and of the role of observables; the next section describes the principle of superposition of states. We then discuss the Lorentz transformations and introduce the concepts of invariance and conservation; the chapter finishes with a section on the distinction between laboratory and centre-of-mass coordinates.

1.2 The Postulates of Quantum Mechanics

In general, we use the Schrödinger method; we shall start by considering the Schrödinger equation and its interpretation. We assume the existence of a function $\psi(x, t)$ which is a function of position x, and of time t. This function is often called a wave function since the earliest types studied were analytically the same as classical waves; however, a more correct term is "state function" since it describes the state of a particle. We shall use the term, wave function, or, if necessary to stress a point, state function. Similar functions which describe the state of an assembly of particles will have more independent variables of position.

$\psi(x, t)$ for a single particle is assumed to satisfy the Schrödinger equation

$$- \frac{\hbar^2}{2m} \nabla^2 \psi(x, t) + V(x) \psi(x, t) = + i\hbar \frac{\partial}{\partial t} \psi(x, t), \qquad (1.1)$$

where m is the mass of the particle, $\hbar$ is Planck's constant divided by 2π, and $V(x)$ is the potential energy of the particle at x.

If the energy E of the particle is a constant, we have

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x}, t) = E\psi(\mathbf{x}, t). \tag{1.2}$$

A solution of Eq. (1.1) is

$$\psi(\mathbf{x}, t) = \psi(\mathbf{x}) \exp(-iEt/\hbar),$$

where $\psi(\mathbf{x})$ satisfies

$$H\psi = E\psi. \tag{1.3}$$

H represents $(-\hbar^2/2m)\nabla^2 + V(\mathbf{x})$ and is called the Hamiltonian operator.

The important properties of a physical system are the quantities which can be measured or observed; such quantities are called observables and any theory which is to predict the behaviour of the system must give the values of these observables. If the system is described by a state function, then there must be means of predicting the observables from this function and the procedure for doing this is given by a set of postulates. These postulates are the rules for interpreting the state function $\psi(\mathbf{x}, t)$. They are:

(1) To every observable there corresponds an operator.

(2) The result of one measurement of an observable is one of the eigenvalues of the equation

$$P\phi = p\phi,$$

where P is the operator and p is an eigenvalue.

(3) The average value of a large number of measurements of an observable is given by the "expectation value"

$$\langle P \rangle = \int \psi^* P\psi \, d\mathbf{x}, \tag{1.4}$$

providing that $\int \psi^* \psi \, dx = 1$ and that there exist suitable boundary conditions. (The integration is performed over all the space within the boundary. The asterisk indicates the complex conjugate.)

In future we shall abbreviate the integration $\int \psi^* P\psi \, dx$ to $(\psi | P | \psi)$. In general, the "matrix element" $(\phi | P | \psi) = \int \phi^* P\psi \, dx$. The integral is not always over space $(\int dx)$: it may be necessary to imply integration over other continuous or discontinuous variables.

The formal properties of operators are important and we state some of these.

(a) They are linear: that is, if $\psi = \sum_n c_n \psi_n$, we have that

$$P\psi = \sum_n c_n P\psi_n.$$

(b) They obey the laws of association and distribution: that is, if P, Q, and R are operators, we have that

$$P(QR) = (PQ)R \quad \text{and} \quad P(Q+R) = PQ + PR.$$

(c) They are Hermitian: that is,

$$\int \phi^* P \psi \, d\mathbf{x} = \int \psi P^* \phi^* \, d\mathbf{x}.$$

(d) As a result of (c) their eigenvalues are real:

$$P\psi_n = p_n \psi_n \quad \text{and} \quad P^* \psi_n^* = p_n \psi_n^*.$$

(e) In postulate 2 the eigenvalues may take on continuous or discrete values depending on the physical situation. Our notation represents the discrete set of eigenfunctions of the operator P by

$$\psi_1, \psi_2, \ldots, \psi_n, \ldots$$

which have the eigenvalues $p_1, p_2, \ldots, p_n, \ldots$, respectively. Then the eigenfunctions form an orthogonal set

$$\int \psi_n^* \psi_m \, dx = (\psi_n | \psi_m) = 0, \quad m \neq n.$$

The eigenfunctions are orthonormal if

$$(\psi_n | \psi_m) = \delta_{nm}.$$

(f) If the state function is an eigenfunction of the operator P, then a measurement of the corresponding observable can only yield a value equal to the eigenvalue.

(g) Operators do not necessarily obey a commutative law; that is, two operators P and Q need not have $PQ\psi = QP\psi$. If $PQ\psi = QP\psi$, the operators are said to commute and this statement is written $[P,Q] = 0$. The state function is implied thus: $[P,Q]\psi = (PQ - QP)\psi$ for all ψ.

(h) If $[P,Q] = 0$, then $[P,f(Q)] = 0$, where $f(Q)$ is a function of Q which can be expanded as a power series in Q.

(i) If $[P,Q] = 0$ and $P\psi_n = p_n \psi_n$, then $(\psi_m | Q | \psi_n) = 0$ if $m \neq n$.

(j) If $[P,Q] = 0$, then it is possible to find a set of functions which are simultaneously eigenfunctions of P and Q. If $[P,Q] \neq 0$, this cannot be done except for a state ψ which has $[P,Q]\psi = 0$.

(k) If $[P,Q] = 0$ and $[P,R] = 0$ but $[Q,R] \neq 0$, then it is not possible to find functions which are simultaneously eigenfunctions of P, Q,

and R. It is only possible to find eigenfunctions for P and Q *or* for P and R. To correspond to this we can only make simultaneous measurements of the observables corresponding to the pair of operators P and Q *or* to the pair P and R: it is not possible to measure all three together. The method of measurement determines the pair observed. The exception occurs for functions which have a zero eigenvalue for Q (or R): these can also be eigenfunctions of R (or Q). It follows that we can make simultaneous measurements of all three observables on a state represented by such eigenfunctions. Similar properties are obvious for the case of more than three operators.

(l) If $\psi_1, \psi_2, ..., \psi_n, ...$ are an orthonormal set but are not eigen-functions of P and Q, then

$$(\psi_m | PQ | \psi_n) = \sum_r (\psi_m | P | \psi_r)(\psi_r | Q | \psi_n).$$

The postulates 1, 2, and 3 form the basis of the interpretation of quantum mechanics and the properties of the required operators are the rules by which the interpretation is conducted. In the Schrödinger method the state function is the important part of the mechanics; however, it is also possible to develop certain aspects not by referring to the actual form of the state function, but by developing the formal properties of operators. We shall use this technique in the chapter on angular momentum.

We must end this section by briefly clarifying the methods of manipulating operators and the notation we shall employ. Firstly, we define the adjoint or Hermitian conjugate ($P^\dagger$) of an operator P by the equation

$$(P^\dagger \psi | \phi) = (\psi | P\phi).$$

An operator which is Hermitian has $P^\dagger = P$. It is now easy to prove result (d). Suppose

$$P\psi_n = p_n \psi_n.$$

Then we have

$$(\psi_n | P\psi_n) = p_n(\psi_n | \psi_n).$$

But from the definition of an Hermitian conjugate and of an Hermitian operator we have

$$(\psi_n | P\psi_n) = (P^\dagger \psi_n | \psi_n) = (P\psi_n | \psi_n).$$

Now the matrix elements have the property that

$$(\phi | \psi)^* = (\psi | \phi).$$

Hence we have

$$(\psi_n | P\psi_n) = (\psi_n | P\psi_n)^*.$$

Since we require $(\psi_n | \psi_n)$ to be real, it follows that the eigenvalues of P are real. Finally, we note that a unitary operator U has the property $U^\dagger = U^{-1}$ where $UU^{-1} = 1$.

1.3 The Principle of the Superposition of States

A further important principle in quantum mechanics is that of the superposition of states; we shall discuss the principle briefly in order to indicate where its use is essential to describe a physical process and to make the reader aware that it is in almost continual use. For a more complete discussion the reader is referred to the first chapter of Dirac's "Quantum Mechanics" (1947).

Suppose a physical system is in a state which is one of a finite or infinite orthonormal set, $viz.$ (e) in Section 1.2. This set of functions ψ, however, may not be a unique way of describing the possible states of the system: suppose there exist a second set of functions ϕ, which also form an orthonormal set and are not in one to one correspondence with the set ψ. We shall have that the ϕ are eigenvalues of an operator Q with corresponding observable Q', whilst ψ are eigenfunctions of P with observable P'. Then a measurement of Q' will always yield an eigenvalue of the equation

$$Q\phi = q\phi.$$

Now ψ is not an eigenfunction Q; however, the principle of superposition effectively means that ψ can be expanded as a linear sum of the eigenfunctions ϕ. The orthonormality of the functions permits this analytically but the principle has deeper significance than reinforcing that possibility. Thus

$$\psi_n = \sum_m a_{nm} \phi_m. \tag{1.5}$$

Then if we make repeated measurement of Q' the average value observed $\langle Q \rangle_n$ for the state ψ_n is given by applying postulate 3, Section 1.2,

$$\langle Q \rangle_n = \int \psi_n^* Q \psi_n \, dx.$$

Substitute Eq. (1.5), then

$$\langle Q \rangle_n = \int \left(\sum_m a_{nm}^* \phi_m^* \right) Q \left(\sum_l a_{nl} \phi_l \right) dx.$$

Using the orthonormality of the set ϕ we have

$$\langle Q \rangle_n = \sum_m a_{nm}^* a_{nm} q_m. \tag{1.6}$$

Equation (1.6) also effectively states that if the system is in state ψ_n, then the probability of finding it in state ϕ_m, by measurement of Q',

is $(a_{nm}^* a_{nm})$. Thus the observer who makes repeated measurements of Q' on the state ψ_n sees the states $\phi_1, \phi_2, \phi_3, ..., \phi_m, ...$ with respective weights $(a_{n1}^* a_{n1}), (a_{n2}^* a_{n2}), (a_{n3}^* a_{n3}), ..., (a_{nm}^* a_{nm}), ...$ The sum of these intensities must be unity.

The state ψ_n therefore appears to be a linear superposition of the states $\phi_1, \phi_2, \phi_3, ..., \phi_m, ...$ The inverse is also true: a state ϕ_l will appear to be a linear superposition of the states $\psi_1, \psi_2, ..., \psi_n, ...$ if a measurement is made of P'. The fact that a particular state can appear to be the sum of other states is a necessary assumption in quantum mechanics; the assumption is defined by the principle of superposition of states.

We can make an extension to this treatment by formally omitting the act of making a measurement of the observables P' or Q'. The probability of finding a certain result to the measurement becomes the probability of finding the system in the state which is described by the eigenfunction having the result as eigenvalue. Thus the probability of finding a system described by the state function ψ_n in a state described by ϕ_m is $(a_{nm}^* a_{nm})$. Now it is easy to prove that

$$| (\phi_m | \psi_n) |^2 = a_{nm}^* a_{nm},$$

and this leads us to generalize: the probability that a system described by a state function ψ is in a state ϕ is

$$| (\phi | \psi) |^2. \tag{1.7}$$

We can mention an excellent example of the meaningfulness of the principle of superposition by introducing the description of polarized light. At the moment we consider only fully polarized light; the complete description of all states of polarization must await Chapter VIII. Classically we know that there are two possible descriptions: one describes a polarized beam by a sum with correct amplitudes and phase of two orthogonal plane polarized beams; the other describes the beam by a sum, with correct amplitudes and phase, of two opposite circularly polarized beams. This kind of summation is exactly like that of Eq. (1.5), and it is possible to construct state functions for all the states of polarization and to find the correct linear sum in either description which describes the state of polarization of photons in the beam. Similarly, operators can be defined which will predict the probability of finding a certain result if a particular measurement of the state of polarization is made on the beam of photons: for example, that a measurement of the state of circular polarization of a plane-polarized photon finds the photon having left- or right-hand circular

polarization with equal probability. This description does not add anything to our knowledge of polarized light but does simplify certain problems.

So far in this section the language used has been simple; it is usual to call the set of functions ψ or ϕ base vectors and the space in which they exist Hilbert space. A rotation of the axes of this Hilbert space transforms the components of the stationary vector; this is equivalent to transforming from one description to the other, as is implied in Eq. (1.5), that is, from one set of base vectors to another set. In fact, analytically the connection between the components of a vector in rotated and in unrotated space is similar to Eq. (1.5). We shall often use superposition without mentioning it; the vector addition of angular momentum and the partial wave analysis are two examples.

1.4 The Lorentz Transformations

All modern quantum mechanics can only proceed meaningfully from its assumptions if these satisfy the requirements of the special theory of relativity (the only exception is the original Schrödinger wave equation; for this reason it can only be used in restricted cases). We will now discuss those postulates of relativity which will affect our subsequent material.

The special theory of relativity postulates that:

(1) The laws of physics formulated by an observer are independent of the state of uniform motion or position of the observer and his apparatus.

(2) The velocity of light *in vacuo* is the same for all such observers.

Let us see how this works by considering two observers who have agreed to make identical observations on identical phenomena. We must distinguish carefully between the adjectives "same" and "identical". We take the words "same phenomenon" to mean a phenomenon which has a common source for both observers; for example, the diffraction of light from a specified star. The same words "same observation" would mean the measurement of a quantity associated with the common phenomenon; for example, the wavelength of a particular hydrogen spectral line in the light from our specified star. Obviously the same quantities will depend on the relative motion of our observers: in our example the Doppler effect alters the wavelength. By "identical phenomena" we mean phenomena which are alike in all respects but do not have a common source; the "identical observations" are the separate measurements of a quantity associated with both these phenomena. Two observers who separately measure the wavelength of a cadmium spectral line from their own local lamp

would be measuring identical quantities but not the same quantity. As long as the measurements are not performed in a gravitational field or under conditions of acceleration, two such observers of identical quantities must yield identical results which will be independent of the two observers' state of relative motion. It follows that if our two observers set up separate equations of motion which describe the observed phenomena the two equations will be formally identical.

The system of coordinates which each observer sets up (normally stationary to himself) is called an inertial frame. The coordinate axes of two inertial frames can be connected by one or more of the following: rotation about the origin, displacement of one origin from the other, uniform relative velocity. As observers we are restricted to one inertial frame, but we do know how to transform observables and equations to the values and forms they would have if observed from another inertial frame (here we mean an observable quantity which is the same in the sense defined earlier). This system of transformations is called the Lorentz transformations, after their originator. We have already stated that in formulating a theory for any physical phenomenon it is essential that the transformation of our equations to the inertial frame of a second observer does not change the consequence of the theory when this requirement is satisfied the theory is said to be Lorentz covariant. Maxwell's electromagnetic theory, Dirac's equation for the electron, and the Klein–Gordon equation for bosons are all Lorentz covariant. Schrödinger's equation is not.

We shall now consider the Lorentz transformations. We commence by defining two kinds of transformations:

(1) A "passive or coordinate transformation" is one in which the physical system is unchanged and we consider the relation between a quantity as observed from one inertial frame and the same quantity as observed from a second inertial frame; geometrically, this corresponds to a displacement and rotation of a four-dimensional coordinate system.

(2) In contrast, an active transformation is one in which a Lorentz transformation is applied to the physical system while the observer's inertial frame is unchanged; geometrically, this corresponds to a rotation and displacement of the physical system within a fixed four-dimensional coordinate system.

In mathematical formalism these two types of transformation are indistinguishable; however, for clarity in discussion we shall consider only coordinate transformations. We are therefore interested in the relation between the results of measurements made by two separate observers on the same quantity.