The Physical Interpretation of the Quantum Dynamics.

By P. A. M. Dirac, St. John's College, Cambridge; Institute for Theoretical Physics, Copenhagen.

(Communicated by R. H. Fowler, F.R.S.—Received December 2, 1926.)

§ 1. *Introduction and Summary.*

The new quantum mechanics consists of a scheme of equations which are very closely analogous to the equations of classical mechanics, with the fundamental difference that the dynamical variables do not obey the commutative law of multiplication, but satisfy instead the well-known quantum conditions. It follows that one cannot suppose the dynamical variables to be ordinary numbers (c-numbers), but may call them numbers of a special type (q-numbers). The theory shows that these q-numbers can in general be represented by matrices whose elements are c-numbers (functions of a time parameter).

When one has performed the calculations with the q-numbers and obtained all the matrices one wants, the question arises how one is to get physical results from the theory, *i.e.*, how can one obtain c-numbers from the theory that one can compare with experimental values? Hitherto this has been done with the help of a number of special assumptions. In Heisenberg's original matrix mechanics it was assumed that the elements of the diagonal matrix that represents the energy are the energy levels of the system, and the elements of the matrix that represents the total polarisation, which are periodic functions of the time, determine the frequencies and intensities of the spectral lines in analogy to the classical theory. Schrödinger's wave representation of the quantum mechanics has provided new ways of obtaining physical results from the theory, based on the assumption that the square of the amplitude of the wave function can in certain cases be interpreted as a probability. From this assumption one can, for instance, work out the probability of a transition being produced in a system (or the number of transitions produced in an assembly of like systems) by an arbitrary external perturbing force,* and can thus, by supposing the perturbation to consist of incident radiation, obtain directly Einstein's B coefficients. Again in Born's treatment of collision problems† it is assumed that the square of the amplitude of the wave function scattered

* See Schrödinger, 'Ann. d. Phys.,' vol. 81, p. 112 (1926); also § 5 of the author's paper 'Roy. Soc. Proc.,' A, vol. 112, p. 661 (1926).

† Born, 'Z. f. Physik,' vol. 37, p. 863; vol. 38, p. 803 (1926).

in any direction determines the probability of the colliding electron (or other body) being scattered in that direction.

Recently Heisenberg has obtained another point of contact between the theory and experiment, of a somewhat different nature.* If one considers the problem of two atomic systems in resonance, *i.e.*, with energy pulsating from one to the other, one can find the time mean of the energy of one of them, by assuming this time mean to be given by a diagonal element of the matrix that represents the energy of that system. Similarly one can find the time mean of the square of its energy, and of the cube of its energy, and so on. Heisenberg has shown that these calculated time means are just what one would expect from the assumption that the energy changes discontinuously from one quantised value to another. The theory can thus be considered to show that the energy actually does change discontinuously from one quantised value to another, and it enables one to calculate the fraction of the total time during which the energy has any particular value, but it can give no information about the times of the transitions.

This result is capable of wide extensions. It can be applied to any dynamical system, not necessarily one composed of two parts in resonance with one another, and to any dynamical variable, not necessarily one that can take only quantised values. One can (disregarding difficulties introduced by degeneration) calculate the time mean of any dynamical variable, g say, for each stationary state of the system, and similarly the time mean of g^2, and of g^3, etc. The information thus obtained about g regarded as a function of the time can be summed up by one stating the fraction of the total time during which g lies between any two specified numerical values, g' and g'' say. One can say nothing about the intervals of time during which this condition is satisfied except the fraction they form of the whole time.

It thus appears that certain questions that one can ask about the system on the classical theory (*e.g.*, the question : For what fraction of the total time does g lie between two specified values ?) can be given definite unambiguous answers on the quantum theory as well as on the classical theory. In the present paper a general theory of such questions and the way the answers are to be obtained will be worked out. This will show all the physical information that one can hope to get from the quantum dynamics, and will provide a general method for obtaining it, which can replace all the special assumptions previously used, and perhaps go further. The questions considered above concerning the fraction of the total time during which specified conditions hold do not form a

* I am indebted to Dr. Heisenberg for informing me of his results before publication.

suitable starting point for this investigation, because they can be given definite answers only for non-degenerate systems, and a system is always degenerate when two or more of its first integrals can take continuous ranges of values. We shall therefore approach the subject from a more general point of view.

The general question of classical mechanics can be formulated as follows: What is the value of any constant of integration* g of a given dynamical system for any given initial conditions, specified by numerical values q_{r0}', p_{r0}', say, for the initial co-ordinates and momenta q_{r0}, p_{r0}? The dynamical theory enables one to express g as a function of the q_{r0}, p_{r0}, and one has then only to substitute for the q_{r0}, p_{r0} the numerical values q_{r0}', p_{r0}' to obtain the answer to the question. On the quantum theory one can also obtain an expression for g as a function of the q_{r0}, p_{r0}, but the q_{r0} and p_{r0} do not now satisfy the commutative law of multiplication, so that if one substituted numerical values for them the result would in general depend on the order in which they were previously arranged. One can thus give no unambiguous answer to the question on the quantum theory.

One cannot answer any question on the quantum theory which refers to numerical values for both the q_{r0} and the p_{r0}. One would expect, however, to be able to answer questions in which only the q_{r0} or only the p_{r0} are given numerical values, or, more generally, when any set of constants of integration ξ_r that commute with one another are given numerical values. If η_r are the variables canonically conjugate to the ξ_r, one would now want to know what one can find out about g, considered as a function of the η_r, with these numerical values for the ξ_r. It will be shown that one can determine without ambiguity the fraction of the whole of η-space for which g lies between any two specified numerical values. More generally, if g_1, g_2 . . . are a set of constants of integration that commute with one another, one can determine the fraction of the whole of η-space for which each g_r lies between specified numerical values. Hence if one is given an assembly of like systems all having the same numerical values for the ξ_r, and one assumes that they are distributed uniformly over the η-space, one can determine the number of systems having each of their g_r's lying between specified numerical values. Questions of this type appear to be the only ones to which the quantum theory can give a definite answer, and they are probably the only ones to which the physicist requires an answer.

To answer questions in which the ξ_r are given numerical values, we require

* The words constant of integration include such quantities as the value of a varying co-ordinate or momentum at a specified time $t = t_0$. On the quantum theory such a "value" would be a q-number, t_0 being, of course, a c-number.

a scheme of matrices to represent the dynamical variables, whose rows and columns refer to numerical values for the ξ_r. In most atomic problems the electrons are given to be initially in definite orbits. For such problems one would take the ξ_r to be the initial values of the action variables J_r (or other first integrals that define the orbits), and could then, with the help of the ordinary matrix representation, work out the fraction of the w-space for which certain specified conditions hold. There are certain problems, however, for which the electrons are not initially in definite orbits (*e.g.*, the problem of the interaction of a β-particle emitted by a radio-active atom with the orbital electrons of the atom, for which the β-particle is initially in the nucleus). To treat such problems we should require a matrix representation of the dynamical variables whose rows and columns refer to other constants of integration of the system than the action variables, such that the initial conditions can be stated by specifying numerical values for these constants of integration. (In the example of the β-particle it would probably be convenient to have the matrix rows and columns referring to the co-ordinates of the β-particle at the time of emission, $t = t_0$, say. We should then be interested only in those rows and columns of the matrices that refer to the β-particle being in the nucleus at the time $t = t_0$, and could calculate the range in which its initial momentum must lie in order that any special kind of interaction, specified by numerical values for certain constants of integration, may take place. We should thus obtain the probability of that kind of interaction, on the assumption that all directions of emission are equally probable.)

We therefore require a theory of the more general schemes of matrix representation, in which the rows and columns refer to any set of constants of integration that commute, and of the laws of transformation from one such scheme to another. This is worked out in §§ 3–5. This theory may be regarded as a development of Lanczos's field theory,* the field representation in Lanzcos's theory being really the same as a matrix representation with matrices that have continuous ranges of rows and columns instead of the usual discrete sets.

In § 6 the transformation theory is used in the investigation of the general method of obtaining physical results from the matrix mechanics, and in § 7 it is shown that this general method is in agreement with the special assumptions previously used.

§ 2. *Notation.*

In the ordinary matrix mechanics one obtains matrices to represent the dynamical variables whose rows and columns refer to stationary states of the

* Lanczos, 'Zeits. f. Phys.,' vol. 35, p. 812 (1926).

system. Thus if $\alpha_1, \alpha_2 \ldots \alpha_u$ are the first integrals of the equations of motion (action variables or otherwise) u being the number of degrees of freedom, each row or column can be labelled by specified values for $\alpha_1, \alpha_2 \ldots \alpha_u$, say $\alpha_1', \alpha_2' \ldots \alpha_u'$, and we may write the elements of the matrix representing any dynamical variable g by $g(\alpha_1', \alpha_2' \ldots \alpha_u'; \alpha_1'', \alpha_2'' \ldots \alpha'')$ or by $g(\alpha' \alpha'')$ for brevity. These matrix elements are functions of the time only. In the present paper we shall not take relativity mechanics into account, and shall count the time variable wherever it occurs as merely a parameter (a c-number).

The parameters that label the rows and columns of the matrices may take either discrete sets of values or all values in certain continuous ranges, or perhaps both. It would complicate the formulæ unnecessarily if we were to write them so as to take both possibilities into account. The case with the continuous ranges of values is the more general and typical one. We shall therefore write all our formulæ as though these parameters can take only continuous ranges of values, it being understood that the necessary changes have to be made when the discrete sets occur. The matrix law of multiplication will now read

$$ab(\alpha' \alpha'') = \int a(\alpha' \alpha''') \, d\alpha''' \, . \, b(\alpha''' \alpha''),$$

where $d\alpha'''$ means $d\alpha_1''' \, . \, d\alpha_2''' \ldots d\alpha_u'''$, and the range of integration is over all the values of the α''''s that label rows and columns of the matrices.*

One cannot go far in the development of the theory of matrices with continuous ranges of rows and columns without needing a notation for that function of a c-number x that is equal to zero except when x is very small, and whose integral through a range that contains the point $x = 0$ is equal to unity. We shall use the symbol $\delta(x)$ to denote this function, *i.e.*, $\delta(x)$ is defined by

$$\delta(x) = 0 \quad \text{when} \quad x \neq 0,$$

and

$$\int_{-\infty}^{\infty} x \, \delta(x) = 1.$$

Strictly, of course, $\delta(x)$ is not a proper function of x, but can be regarded only as a limit of a certain sequence of functions. All the same one can use $\delta(x)$ as though it were a proper function for practically all the purposes of quantum mechanics without getting incorrect results. One can also use the differential coefficients of $\delta(x)$, namely $\delta'(x)$, $\delta''(x) \ldots$, which are even more discontinuous and less " proper " than $\delta(x)$ itself.

* Whenever the limits of integration of an integral are not specified, the whole range of the parameter that is used for labelling matrix rows and columns is to be understood.

A few elementary properties of these functions will now be given so as not to interrupt the argument later. We can obviously take $\delta(-x) = \delta(x)$, $\delta'(-x) = -\delta'(x)$, etc. The condition $\delta(x) = 0$ except when $x = 0$ may be expressed by the algebraic equation $x\,\delta(x) = 0$. [This equation, together with the equation $\delta(x) \cdot x = 0$, can be used in the definition of $\delta(x)$ when x is a q-number or matrix.] If $f(x)$ is any regular function of x and a is any c-number, we have

$$\int_{-\infty}^{\infty} f(x)\,\delta(a-x)\,dx = f(a), \tag{1}$$

so that the operation of multiplying by $\delta(a-x)$ and integrating with respect to x is equivalent to the operation of substituting a for x. Again we have, by integration by parts

$$\int_{-\infty}^{\infty} f(x)\,\delta'(a-x)\,dx = \Big[-f(x)\,\delta(a-x)\Big]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} f'(x)\,\delta(a-x)\,dx = f'(a), \tag{1'}$$

since the integrated term vanishes at both limits, and in general

$$\int_{-\infty}^{\infty} f(x)\,\delta^{(n)}(a-x)\,dx = f^{(n)}(a), \tag{1''}$$

so that the operation of multiplying by $\delta^{(n)}(a-x)$ and integrating with respect to x is equivalent to the operation of differentiating n times with respect to x and susbtituing a for x.

We shall now show that, if b is another c-number,

$$\int_{-\infty}^{\infty} \delta(a-x)\,\delta(x-b)\,dx = \delta(a-b). \tag{2}$$

If we regard the left-hand side as a function of b and put it equal to $\phi(b)$, then we shall have that $\phi(b)$ is equal to zero when b differs appreciably from a, and also

$$\int_{-\infty}^{\infty} \phi(b)\,db = \int_{-\infty}^{\infty} \delta(a-x)\,dx \int_{-\infty}^{\infty} \delta(x-b)\,db = 1.$$

Hence $\phi(b)$ has all the properties of $\delta(a-b)$ and may be put equal to $\delta(a-b)$. We should have obtained equation (2) if we had put $f(x)$ equal to $\delta(x-b)$ in (1). This is thus a case in which one can use $\delta(x-b)$ as though it were a regular function of x without getting a wrong result. Another such case is the putting

of $f(x)$ equal to $\delta(x-b)$, or, more generally, $\delta^{(n)}(x-b)$, in (1′), leading to the equations

$$\int_{-\infty}^{\infty} \delta'(a-x)\,\delta(x-b)\,dx = \delta'(a-b), \tag{2′}$$

and

$$\int_{-\infty}^{\infty} \delta'(a-x)\,\delta^{(n)}(x-b)\,dx = \delta^{(n+1)}(a-b). \tag{2″}$$

Equation (2′) can be verified independently by the partial differentiation of (2) with respect to a, and then (2″) can be verified by the partial differentiation of (2′) n times with respect to b.

From (1′) if we take $f(x) = x$ and $a = 0$, we get

$$\int_{-\infty}^{\infty} -x\,\delta'(x)\,dx = 1.$$

But $-x\,\delta'(x)$, considered as a function of x, vanishes except when $|x|$ is very small. Hence $-x\,\delta'(x)$ has all the properties of $\delta(x)$, and we can write

$$-x\,\delta'(x) = \delta(x). \tag{3}$$

When using matrices with continuous ranges of rows and columns, we need the function $\delta(x)$ to express the elements of the unit matrix. The unit matrix must by definition be such that when multiplied by any matrix y the product is equal to that matrix, *i.e.*, we must have

$$\int 1(\alpha'\alpha''')\,d\alpha'''\,.\,y(\alpha'''\alpha'') = y(\alpha'\alpha'').$$

We therefore see that

$$1(\alpha'\alpha''') = \delta(\alpha_1' - \alpha_1''')\,.\,\delta(\alpha_2' - \alpha_2''')\ldots\delta(\alpha_u' - \alpha_u''') = \delta(\alpha' - \alpha'''),$$

say, for brevity. The general diagonal matrix $f(\alpha)$ has elements $f(\alpha')\,.\,\delta(\alpha' - \alpha'')$. We shall call the quantities $f(\alpha')$ the diagonal elements of this matrix.

§ 3. *The Transformation Equations.*

The solving of a problem in Heisenberg's matrix mechanics consists in finding a scheme of matrices to represent the dynamical variables, satisfying the following conditions :—

(i) The quantum conditions, $q_r p_r - p_r q_r = ih$, etc.

(ii) The equations of motion, $g\mathrm{H} - \mathrm{H}g = ih\dot{g}$, or if g involves the time explicitly $g\mathrm{H} - \mathrm{H}g + ih\,\partial g/\partial t = ih\dot{g}$.

(iii) The matrix representing the Hamiltonian H must be a diagonal matrix.

(iv) The matrices representing real variables must be Hermitian.

The scheme of matrices that satisfies these conditions is not, in general, unique. If to each of the matrices, g say, we apply the canonical transformation

$$G = bgb^{-1}, \tag{4}$$

where b is any matrix, the new matrices G will satisfy all the algebraic relations that the original ones did ; in particular they will satisfy the quantum conditions. Also, if the elements of the matrix b are not functions of the time, so that we have $\dot{G} = b\dot{g}b^{-1}$, the new matrices will satisfy the equations of motion. Further, if b commutes with H, the new matrix representing the Hamiltonian will be a diagonal matrix, and if in addition the elements of the matrices b and b^{-1} satisfy the condition that $b(\alpha'\alpha'')$ and $b^{-1}(\alpha''\alpha')$ are conjugate imaginaries, each matrix G will be Hermitian when the corresponding matrix g is Hermitian. Thus when these conditions are satisfied the new matrices will satisfy conditions (i) to (iv), and will be just as good as the original ones for representing the dynamical variables. We shall work out the theory of these transformations, and also of the more general kind of transformation to a scheme of matrices that need satisfy only conditions (i) and (ii), which means that b and b^{-1} need satisfy only the conditions that their matrix elements do not involve the time t.

Equation (4) may be written

$$G(\alpha'\alpha'') = \int\int b(\alpha'\alpha''')\, d\alpha''' \,.\, g(\alpha'''\alpha^{(4)})\, d\alpha^{(4)} \,.\, b^{-1}(\alpha^{(4)}\alpha''). \tag{5}$$

When we make a transformation of this kind, we can at the same time make any permutation of the rows of the new matrices G and the same permutation of their columns, without interfering with any of the conditions (i) . . . (iv) that they satisfy. *There is thus no one-one correspondence between the rows and columns of the new matrices and those of the original matrices.* The notation used in equation (5) is unsatisfactory because it implies that there is such a one-one correspondence, the same label α' or $(\alpha_1'\alpha_2' \ldots \alpha_u')$ being used to specify a row and column both of the matrices G and of the matrices g. We therefore modify the notation and write equation (5) thus :—

$$G(\xi_1'\xi_2' \ldots \xi_u' ; \quad \xi_1''\xi_2'' \ldots \xi_u'') = G(\xi'\xi'') = \int\int b(\xi'\alpha')\, d\alpha' \,.\, g(\alpha'\alpha'')\, d\alpha'' \,.\, b^{-1}(\alpha''\xi''), \tag{5'}$$

where the new parameters ξ_r' are quite unconnected with the α''s. The ξ''s may, in fact, take quite different ranges of values from the α''s, or one could even have the ξ''s taking only discrete sets of values while the α''s can take continuous ranges of values, or *vice versa*.

The question now arises as to how one is to label the rows and columns of the new matrices G, *i.e.*, how is one to assign to each row and corresponding column a set of numerical values for the parameters ξ_r'. To do this in a reasonable manner one must find those functions of the dynamical variables, $\xi_1, \xi_2 \ldots \xi_u$, say, that are diagonal matrices in the new scheme of matrix representation, and then assign to each row and corresponding column the value ξ_r' of the diagonal element lying in that row and column of each ξ_r. The labelling is thus carried out so as to make the ξ_r have the matrix elements

$$\xi_r(\xi'\xi'') = \xi_r' \,\delta(\xi_1' - \xi_1'') \,.\, \delta(\xi_2' - \xi_2'') \ldots \delta(\xi_u' - \xi_u'') = \xi_r' \,\delta(\xi' - \xi''), \quad (6)$$

say. The dynamical variables ξ_r are used in the labelling of the rows and columns of the new matrices in exactly the same way in which the dynamical variable α_r were used in the labelling of the rows and columns of the original matrices.

The ξ_r must be constants of integration of the system, since their matrix elements do not involve t. Also they must commute with one another, since diagonal matrices always commute. The ξ_r therefore form a set of canonical co-ordinates, and they will have a set of canonically conjugate momenta $\eta_1, \eta_2 \ldots \eta_u$, say.

The matrices b and b^{-1} satisfy the relations $bb^{-1} = 1$ and $b^{-1}b = 1$, or

$$\int b(\xi'\alpha') \, d\alpha' \,.\, b^{-1}(\alpha'\xi'') = \delta(\xi' - \xi'')$$

and

$$\int b^{-1}(\alpha'\xi') \, d\xi' \,.\, b(\xi'\alpha'') = \delta(\alpha' - \alpha'').$$

The matrix elements $b(\xi'\alpha')$ and $b^{-1}(\alpha'\xi')$ thus form two mutually orthogonal and normalised systems of functions, either if they are regarded as functions of the α''s specified by values for the parameters ξ', or if they are regarded as functions of the ξ''s specified by values for the parameters α'. Any two such mutually orthogonal and normalised systems of functions define a transformation to a new scheme of matrices that satisfies conditions (i) and (ii). If in addition $b(\xi'\alpha')$ and $b^{-1}(\alpha'\xi')$ are conjugate imaginaries, the new scheme of matrices will also satisfy condition (iv). In order that the new scheme may satisfy condition (iii), the ξ's must commute with H. It would follow, since the ξ's are constants of integration, that they must be functions of the original canonical variables q_r and p_r that do not involve t explicitly.

We shall now simplify the notation a little. In equation (5′) there is no need to use the different symbols g and G to denote the same dynamical variable

represented according to the old or the new schemes, since the parameters (ξ' and ξ'' or α' and α'', as the case may be) show quite definitely to which scheme any matrix element belongs. We shall therefore use always the same symbol, g for instance, to denote any particular dynamical variable, and shall write its matrix elements according to the different schemes $g(\alpha'\alpha'')$, $g(\xi'\xi'')$. Further the transformation functions $b(\xi'\alpha')$ and $b^{-1}(\alpha'\xi')$ are sufficiently defined if we write them simply as (ξ'/α') and (α'/ξ'). We can thus write equation (5′) in the simplified form

$$g(\xi'\xi'') = \int\int (\xi'/\alpha')\, d\alpha' \,.\, g(\alpha'\alpha'')\, d\alpha''\, (\alpha''/\xi''). \tag{5''}$$

We shall make it a rule always to use unprimed letters such as g or ξ_r to denote dynamical variables (or q-numbers), and primed or multiply primed letters such as ξ' and α'' to denote parameters, representing matrix rows and columns, which can have specified numerical values, and are c-numbers.

The transformation equation (5″) can equally well be written in any of the forms

$$\int\int (\alpha'/\xi')\, d\xi' \,.\, g(\xi'\xi'')\, d\xi''\, (\xi''/\alpha'') = g(\alpha'\alpha'')$$

say, and

$$\left.\begin{aligned} \int g(\xi'\xi'')\, d\xi''\, (\xi''/\alpha') &= \int (\xi'/\alpha'')\, d\alpha'' \,.\, g(\alpha''\alpha') = g(\xi'\alpha') \\ \int (\alpha'/\xi'')\, d\xi'' \,.\, g(\xi''\xi') &= \int g(\alpha'\alpha'')\, d\alpha''\, (\alpha''/\xi') = g(\alpha'\xi') \end{aligned}\right\} \tag{7}$$

say, corresponding respectively to the matrix equations (written in the old notation)

$$b^{-1}\mathrm{G}b = g, \qquad \mathrm{G}b = bg, \qquad b^{-1}\mathrm{G} = gb^{-1},$$

which follow immediately from (4). The expressions $g(\xi'\alpha')$ and $g(\alpha'\xi')$ introduced in equations (7) may be regarded as the elements of two matrices that represent the dynamical variable g according to two new more general schemes, in which the rows and the columns of the matrices refer to different things. There is now no one-one correspondence between the rows and the columns, so that a diagonal matrix in these new schemes has no meaning. The matrices with elements (ξ'/α') and (α'/ξ') are the unit matrices in their respective schemes, since equations (7) show that these matrices multiplied by a matrix representing an arbitrary q-number g give matrices representing g.

If we apply successively two canonical transformations with the matrices b_1 and b_2, *i.e.*,

$$\mathrm{G} = b_1 g b_1^{-1}, \qquad \mathrm{G}^* = b_2 \mathrm{G} b_2^{-1},$$

the result is the same as the single transformation with the matrix b_2b_1, since

$$G^* = b_2b_1gb_1^{-1}b_2^{-1} = (b_2b_1)\ g\ (b_2b_1)^{-1}.$$

Translated into the new notation, this theorem is that if one applies successively two canonical transformations with the transformation functions (ξ'/α'), (α'/ξ') and (κ'/ξ'), (ξ'/κ') respectively, the result is the same as the single transformation with the transformation functions

$$(\kappa'/\alpha') = \int (\kappa'/\xi')\ d\xi'\ (\xi'/\alpha')$$

and

$$(\alpha'/\kappa') = \int (\alpha'/\xi')\ d\xi'\ (\xi'/\kappa').$$

§ 4. *Some Elementary Matrices.*

The matrix elements of the ξ's are given by equation (6). We must now determine the elements of the matrices canonically conjugate to the ξ's It may be shown that the matrices η_r whose elements are defined by

$$\eta_r\ (\xi'\xi'') = -\ ih\ \delta\ (\xi_1' - \xi_1'') \ldots$$
$$\ldots\ \delta\ (\xi_{r-1}' - \xi_{r-1}'')\ .\ \delta'\ (\xi_r' - \xi_r'')\ .\ \delta\ (\xi_{r+1}' - \xi_{r+1}'')\ \ldots\ \delta(\xi_u' - \xi_u''), \quad (8)$$

satisfy the canonical relations

$$\eta_r\eta_s - \eta_s\eta_r = 0, \qquad \xi_r\eta_s - \eta_s\xi_r = 0 \quad (r \neq s)$$

and

$$\xi_r\eta_r - \eta_r\xi_r = ih.$$

The first two of the relations are verified very easily with the help of (2′) and (2). We shall prove the third for the case of a single degree of freedom ($u = 1$). The proof is exactly similar for several degrees of freedom, but is not so easily written down.

For a single degree of freedom we have

$$\xi\ (\xi'\xi'') = \xi'\ \delta\ (\xi' - \xi'')$$

and

$$\eta\ (\xi'\xi'') = -\ ih\ \delta'\ (\xi' - \xi''),$$

so that

$$(\xi\eta - \eta\xi)\ (\xi'\xi'') = -\ ih \int \{\xi'\ \delta\ (\xi' - \xi''')\ .\ \delta'\ (\xi''' - \xi'')$$
$$- \delta'\ (\xi' - \xi''')\ .\ \xi'''\ \delta(\xi''' - \xi'')\}\ d\xi'''$$
$$= -\ ih \int \Big\{\xi'\ \delta\ (\xi' - \xi''')\ .\ \delta'\ (\xi''' - \xi'')$$
$$- \delta\ (\xi' - \xi''')\ \frac{\partial}{\partial \xi'''}[\xi'''\ \delta\ (\xi''' - \xi'')]\Big\}\ d\xi'''.$$

where the second term has been integrated by parts, and hence

$$(\xi\eta - \eta\xi)\,(\xi'\xi'') = -ih\int\{(\xi' - \xi''')\,\delta(\xi' - \xi''')\,\delta'(\xi''' - \xi'') - \delta(\xi' - \xi''')\,\delta(\xi''' - \xi'')\}\,d\xi'''.$$

The first term in the integral vanishes since $(\xi' - \xi''')\,\delta(\xi' - \xi''') = 0$, and the second can be evaluated with the help of (2). We now obtain

$$(\xi\eta - \eta\xi)\,(\xi'\xi'') = ih\,\delta(\xi' - \xi'')$$

as required.

The variables canonically conjugate to the ξ_r are not, of course, uniquely determined when the ξ_r are given, because the variables

$$\eta_r^* = \eta_r + \partial F/\partial\xi_r, \tag{9}$$

where F is any function of the ξ_r, would also be conjugate to the ξ_r when the η_r are. This corresponds to the fact that a matrix representation is not uniquely determined when one is given the variables ξ_r that are diagonal matrices and that label the rows and columns, because one can multiply each row (ξ') by any function $f(\xi')$ of the parameters ξ_r' and divide the corresponding column by the same quantity, as this process does not affect the validity of any matrix equation and does not change the diagonal matrices. This process, though, does change the η_r matrices defined by (8), the new matrices being

$$\eta_r^*(\xi'\xi'') = \eta_r(\xi'\xi'')\frac{f(\xi')}{f(\xi'')} = \eta_r(\xi'\xi'') + \eta_r(\xi'\xi'')\frac{f(\xi') - f(\xi'')}{f(\xi'')}$$
$$= \eta_r(\xi'\xi'') + \frac{\eta_r(\xi'\xi'')}{f(\xi'')}\Sigma_s\frac{\partial f(\xi')}{\partial\xi_s'}(\xi_s' - \xi_s'').$$

Each term in the summation except the term $s = r$ vanishes when multiplied by $\eta_r(\xi'\xi'')$, since the factor $(\xi_s' - \xi_s'')$ vanishes when multiplied by the factor $\delta(\xi_s' - \xi_s'')$ that occurs in $\eta_r(\xi'\xi'')$ when $s \neq r$. On the other hand the factor $(\xi_r' - \xi_r'')$ of the term $s = r$ gets multiplied by the factor $\delta'(\xi_r' - \xi_r'')$ of $\eta_r(\xi'\xi'')$, and the product is, by equation (3), just $-\delta(\xi_r' - \xi_r'')$. We are thus left with

$$\eta_r^*(\xi'\xi'') = \eta_r(\xi'\xi'') + \frac{ih}{f(\xi')}\frac{\partial f(\xi')}{\partial\xi_r'}\delta(\xi' - \xi'')$$
$$= \eta_r(\xi'\xi'') + \frac{ih}{f(\xi)}\frac{\partial f(\xi)}{\partial\xi_r}(\xi'\xi''),$$

which agrees with (9) when one takes $F = ih\log f$.

For the case of a single degree of freedom one finds readily, with the help of (2″), that the matrix elements of η^2 are

$$\eta^2(\xi'\xi'') = (-ih)^2 \int \delta'(\xi' - \xi''')\, d\xi''' \,.\, \delta'(\xi''' - \xi'') = (-ih)^2\, \delta''(\xi' - \xi''),$$

and more generally, by induction, that

$$\eta^n(\xi'\xi'') = (-ih)^n\, \delta^{(n)}(\xi' - \xi'').$$

Hence, if a is an arbitrary c-number, the elements of the matrix $e^{ia\eta}$ are given by

$$e^{ia\eta}(\xi'\xi'') = \Sigma \frac{1}{n!}(ia\eta)^n(\xi'\xi'') = \Sigma \frac{1}{n!}(ah)^n\, \delta^{(n)}(\xi' - \xi'') = \delta(\xi' - \xi'' + ah),$$

with the help of Taylor's expansion theorem.* The matrix $e^{ia\eta}$ thus contains only elements referring to " transitions " in which ξ changes by the amount ah, as is to be expected. Similar results hold for any number of degrees of freedom, but their proofs are not so easily written down.

§ 5. *Transformation Theory.*

We shall now consider the transformation between any two matrix schemes, (ξ) and (α) say, that need satisfy only the conditions (i) and (ii) of § 3. We have

$$\eta_1(\xi'\xi'') = -ih\, \delta'(\xi_1' - \xi_1'')\,.\,\delta(\xi_2' - \xi_2'') \ldots \delta(\xi_u' - \xi_u''),$$

and hence

$$\eta_1(\xi'\alpha') = \int \eta_1(\xi'\xi'')\, d\xi''\, (\xi''/\alpha') = -ih \frac{\partial(\xi'/\alpha')}{\partial \xi_1'},$$

with the help of equations (1) and (1′), and in general

$$\eta_r(\xi'\alpha') = -ih \frac{\partial(\xi'/\alpha')}{\partial \xi_r'}.$$

Again we have

$$\xi_r(\xi'\alpha') = \int \xi_r'\, \delta(\xi' - \xi'')\, d\xi''\, (\xi''/\alpha') = \xi_r'(\xi'/\alpha'),$$

and more generally, if $f(\xi_r)$ is any function of the ξ_r only,

$$f(\xi_r)(\xi'\alpha') = \int f(\xi_r')\, \delta(\xi' - \xi'')\, d\xi''\, (\xi''/\alpha') = f(\xi_r')(\xi'/\alpha').$$

We shall now show that if $f(\xi_r, \eta_r)$ is any function of the ξ_r and their canonical conjugates η_r defined by (8), that is rational and integral in the η_r, then

$$f(\xi_r, \eta_r)(\xi'\alpha') = f\left(\xi_r',\ -ih \frac{\partial}{\partial \xi_r'}\right)(\xi'/\alpha'), \tag{10}$$

* The application of Taylor's theorem to the function $\delta(x)$ appears to be legitimate, because $\delta(x)$ can be regarded as the limit of a sequence of functions for each of which Taylor's theorem holds.

so that the elements of the matrix representing f in the $(\xi'\alpha')$ scheme are given by a certain operator operating on the transformation function (ξ'/α'). It is sufficient to prove that if the theorem is true for any two functions, f_1 and f_2, say, then it is true for their sum $f_1 + f_2$ and their product $f_1 f_2$. The case of the sum is trivial. For the product we have

$$f_1(\xi_r, \eta_r)\, f_2(\xi_r, \eta_r)\, (\xi'\alpha')$$

$$= \iint f_1(\xi_r, \eta_r)\, (\xi'\alpha'')\, d\alpha''\,.\,(\alpha''/\xi''),\, d\xi''\,.\, f_2(\xi_r, \eta_r)\, (\xi''\alpha')$$

$$= \iint f_1\left(\xi_r', -ih\frac{\partial}{\partial \xi_r'}\right)(\xi'/\alpha'')\, d\alpha''\,.\,(\alpha''/\xi'')\, d\xi''\,.\, f_2\left(\xi_r'', -ih\frac{\partial}{\partial \xi_r''}\right)(\xi''/\alpha')$$

$$= f_1\left(\xi_r', -ih\frac{\partial}{\partial \xi_r'}\right)\iint (\xi'/\alpha'')\, d\alpha''\,.\,(\alpha''/\xi'')\, d\xi''\,.\, f_2\left(\xi_r'', -ih\frac{\partial}{\partial \xi_r''}\right)(\xi''/\alpha')$$

$$= f_1\left(\xi_r', -ih\frac{\partial}{\partial \xi_r'}\right) f_2\left(\xi_r', -ih\frac{\partial}{\partial \xi_r'}\right)(\xi'/\alpha'),$$

as required. In the same way it may be shown that

$$f(\xi_r, \eta_r)\, (\alpha'\xi') = f\left(\xi_r', ih\frac{\partial}{\partial \xi_r'}\right)(\alpha'/\xi'). \tag{10'}$$

The formula (10) provides us with a powerful way of obtaining the matrix scheme of representation that makes any specified function of the dynamical variables a diagonal matrix. Suppose, for instance, that we are given a function of the ξ's and η's, $\mathrm{F}(\xi_r, \eta_r)$ say, and we want the matrix scheme, (α) say, that makes F a diagonal matrix, *i.e.*, we want to have

$$\mathrm{F}(\alpha'\alpha'') = \mathrm{F}(\alpha')\,.\,\delta(\alpha' - \alpha''),$$

where $\mathrm{F}(\alpha')$ is some function of the single set of parameters α'. Formula (10) shows that

$$\mathrm{F}\left(\xi_r', -ih\frac{\partial}{\partial \xi_r'}\right)(\xi'/\alpha') = \mathrm{F}(\xi_r, \eta_r)\, (\xi'\alpha') = \int (\xi'/\alpha'')\, d\alpha''\,.\,\mathrm{F}(\alpha''\alpha')$$

$$= \mathrm{F}(\alpha')\, (\xi'/\alpha'). \tag{11}$$

This is an ordinary differential equation for (ξ'/α'), considered as a function of the ξ''s, and its different solutions, when we have found them, are to be specified by different values for the parameters α'. We can then easily obtain the (α) matrix elements of any dynamical variable $f(\xi_r, \eta_r)$ from the formula

$$f(\xi_r, \eta_r)\, (\alpha'\alpha'') = \iint (\alpha'/\xi')\, d\xi'\,.\, f\left(\xi_r', -ih\frac{\partial}{\partial \xi_r'}\right)(\xi'/\alpha'').$$

The characteristic values of the differential equation, denoted by $\mathrm{F}(\alpha')$, are the diagonal elements of the diagonal matrix representing F.

If we take the ξ's and η's to be the ordinary q's and p's of the system at some specified time, and take F to be the Hamiltonian, then equation (11) is just Schrödinger's wave equation, and we get Schrödinger's method of solving a dynamical problem on the quantum theory. *The eigenfunctions of Schrödinger's wave equation are just the transformation functions* (*or the elements of the transformation matrix previously denoted by b*) *that enable one to transform from the* (*q*) *scheme of matrix representation to a scheme in which the Hamiltonian is a diagonal matrix.*

For systems in which the Hamiltonian involves the time explicitly, there will in general be no matrix scheme with respect to which H is a diagonal matrix, since there will be no set of constants of integration that do not involve the time explicitly. For such cases we must find a more general wave equation than equation (11). We shall first show that if $q_{r\tau}$ denotes the value of each q_r at the time $t = \tau$, and if the α's are a set of constants of integration that can be expressed as functions of the q's, p's and t at the arbitrary time t *that do not involve the parameter* τ, then

$$\mathrm{H}\,(q_\tau'\alpha') = ih\frac{\partial}{\partial\tau}(q_\tau'/\alpha').$$

This condition that the α's must satisfy is such that

$$\frac{df}{d\tau}(\alpha'\alpha'') = \frac{\partial}{\partial\tau}[f(\alpha'\alpha'')],$$

where f is any function of the p_τ's and q_τ's. Further, if f does not involve τ explicitly, we must have $ih\,df/d\tau = f\,\mathrm{H}_\tau - \mathrm{H}_\tau f$, where H_τ denotes the Hamiltonian at time τ. Hence

$$\begin{aligned}
&(f\mathrm{H}_\tau - \mathrm{H}_\tau f)\,(q_\tau'\alpha') \\
&\quad = ih\frac{df}{d\tau}(q_\tau'\alpha') = ih\int(q_\tau'/\alpha'')\,d\alpha''\,.\,\frac{df}{d\tau}(\alpha''\alpha') \\
&\quad = ih\frac{\partial}{\partial\tau}\int(q_\tau'/\alpha'')\,d\alpha''\,.\,f(\alpha''\alpha') - ih\int\frac{\partial}{\partial\tau}(q_\tau'/\alpha'')\,.\,d\alpha''\,.\,f(\alpha''\alpha') \\
&\quad = ih\frac{\partial}{\partial\tau}\int f(q_\tau'q_\tau'')\,dq_\tau''\,.\,(q_\tau''/\alpha') - ih\int\frac{\partial}{\partial\tau}(q_\tau'/\alpha'')\,.\,d\alpha''\,.\,f(\alpha''\alpha') \\
&\quad = ih\int f(q_\tau'q_\tau'')\,dq_\tau''\,.\,\frac{\partial}{\partial\tau}(q_\tau''/\alpha') - ih\int\frac{\partial}{\partial\tau}(q_\tau'/\alpha'')\,.\,d\alpha''\,.\,f(\alpha''\alpha'),
\end{aligned}$$

since, when f is a function of the p_τ's and q_τ's that does not involve τ explicitly, $f(q_\tau'\,q_\tau'')$ must be independent of τ. Again, we have

$$(f\mathrm{H}_\tau - \mathrm{H}_\tau f)\,(q_\tau'\alpha') = \int f(q_\tau'q_\tau'')\,dq_\tau''\,.\,\mathrm{H}_\tau(q_\tau''\alpha') - \int\mathrm{H}_\tau(q_\tau'\alpha'')\,d\alpha''\,.\,f(\alpha''\alpha').$$

Comparing these two expressions for $(fH_\tau - H_\tau f)(q_\tau'\alpha')$ we see that, since they hold for any function f, we must have

$$H_\tau (q_\tau'\alpha') = ih\,\partial (q_\tau'/\alpha')/\partial\tau.$$

If we write t for τ and q for q_τ, this becomes

$$H (q'\alpha') = ih\,\partial (q'/\alpha')/\partial t,$$

which can be compared with our previously obtained formula

$$p_r (q'\alpha') = -ih\,\partial (q'/\alpha')/\partial q_r'.$$

We now have

$$H\left(q_r, \quad -ih\frac{\partial}{\partial q_r}\right)(q'/\alpha') = H (q_r, p_r)(q'\alpha') = ih\frac{\partial}{\partial t}(q'/\alpha'), \tag{12}$$

which is Schrödinger's wave equation for Hamiltonians that involve the time explicitly.

The equations of a contact transformation from a set of canonical variables η_r, ξ_r to a set α_r, β_r may, on the classical theory, be put in the simple form

$$\eta_r = \partial S/\partial\xi_r, \qquad \beta_r = \partial S/\partial\alpha_r, \tag{13}$$

where S may be any function of the ξ's and α's. Jordan* has shown that the transformation equations on the quantum theory may also be put in this form provided S is written in the form

$$S = \Sigma f (\xi_r) g (\alpha_r),$$

i.e., all the ξ's in the products occurring in S must be in front of all the α's. (It is to be understood that this order must be preserved when one performs the partial differentiations.) This result follows very easily from the present theory. We have

$$\begin{aligned} f (\xi_r)\, g (\alpha_r)\, (\xi'\alpha') &= \int\int f (\xi_r)\, (\xi'\xi'')\, d\xi''\, (\xi''/\alpha'')\, d\alpha''\, .\, g (\alpha_r)\, (\alpha''\alpha') \\ &= \int\int f (\xi_r')\, \delta (\xi' - \xi'')\, d\xi''\, (\xi''/\alpha'')\, d\alpha''\, .\, g (\alpha_r'')\, \delta (\alpha'' - \alpha') \\ &= f (\xi_r')\, g (\alpha_r')\, .\, (\xi'/\alpha'), \end{aligned}$$

and thus

$$\Sigma f (\xi_r)\, g (\alpha_r)\, (\xi'\alpha') = \Sigma f (\xi_r')\, g\ (\alpha_r')\, .\, (\xi'/\alpha') \tag{14}$$

for any sets of functions $f (\xi_r)$ and $g (\alpha_r)$. Put

$$(\xi'/\alpha') = \exp.\ iS/h,$$

* Jordan, 'Z. f. Physik,' vol. 38, p. 513 (1926).

and suppose S to be written in the form $\Sigma f(\xi') g(\alpha')$. We now have

$$\eta_r(\xi'\alpha') = -ih\frac{\partial}{\partial\xi_r'}(\xi'/\alpha') = \frac{\partial S(\xi', \alpha')}{\partial\xi_r'}.(\xi'/\alpha') = \frac{\partial S(\xi, \alpha)}{\partial\xi_r}(\xi'\alpha')$$

from (14), provided $\partial S(\xi, \alpha)/\partial\xi_r$ is also written in the form $\Sigma f(\xi) g(\alpha)$. Similarly it may be shown that

$$\beta_r(\xi'\alpha') = \frac{\partial S(\xi, \alpha)}{\partial\alpha_r}(\xi'\alpha').$$

These equations are just the equations (13) written as relations between matrix elements.

§ 6. *Physical Interpretation of the Matrices.*

To obtain physical results from the matrix theory, the only assumption one need make is that the diagonal elements of a matrix, whose rows and columns refer to the ξ's say, representing a constant of integration, g say, of the dynamical system, determine the average values of the function $g(\xi_r, \eta_r)$ over the whole of η-space for each particular set of numerical values for the ξ's, in the same way in which they certainly would in the limiting case of large quantum numbers. Thus if

$$g(\xi'\xi'') = g(\xi').\delta(\xi' - \xi'')$$

when the ξ'''s are nearly equal to the ξ''s, we assume that $g(\xi')$ is the average value of g over all η-space when $\xi_r = \xi_r'$. In the case when the diagonal elements of the matrix g are finite without requiring the removal of the factor $\delta(\xi' - \xi'')$, the corresponding assumption is that the diagonal element $g(\xi'\xi')$ is equal to $(2\pi h)^{-1}$ times the integral of $g(\xi_r, \eta_r)$ over all η-space when $\xi_r = \xi_r'$.

This assumption, of course, enables us to determine also the η-average of any function of g, and further, if $g_1, g_2 \ldots g_u$ are a set of g's that commute with one another and are independent functions of the ξ's and η's, we can determine the η-average of any function of these g's. (The g's must commute, since if they did not we should find, for instance, that the η-average of $g_1 g_2$ was different from that of $g_2 g_1$, and we should not know how to interpret these averages physically.) This information appears to be all that one can hope to be able to get concerning the g's as functions of the η's for specified numerical values of the ξ's.

All this information will be summed up if we can determine the fraction of the whole of η-space (or the total volume of η-space, when this fraction is zero) for which each g_r lies between any two given numerical values, *i.e.*, for which $g_r' < g_r < g_r''$, say. To do this we must find the matrix that represents

$$\delta(g_1 - g_1').\delta(g_2 - g_2') \ldots \delta(g_u - g_u') = \delta(g - g'),$$

say, for brevity. If we then integrate this matrix with respect to the parameters g', the result, namely,

$$\int_{g'}^{g''} \delta(g - g')\, dg',$$

will be the matrix which represents that function of the g's that equals unity when $g_r' < g_r < g_r''$, and zero when these conditions are not fulfilled. The diagonal elements of this matrix will then give the η-average of this function, which is just the fraction of the whole of η-space for which $g_r' < g_r < g_r''$ (or else they will give the integral of this function over the whole of η-space, which is just the total volume of η-space for which $g_r' < g_r < g_r''$).

The matrix $\delta(g - g')$ must satisfy the conditions*

$$(g_r - g_r') \cdot \delta(g - g') = \delta(g - g') \cdot (g_r - g_r') = 0$$

for each r, and

$$\int \delta(g - g')\, dg' = 1.$$

It is easily verified that the matrix with elements

$$\delta(g - g')\,(\xi'\xi'') = (\xi'/g') \cdot (g'/\xi'')$$

satisfies these conditions. The last of the conditions is obviously satisfied on account of the orthogonal and normalised properties of the transformation functions (ξ'/g') and (g'/ξ''), while to prove the others we have

$$g_r\, \delta(g - g')\,(\xi'\xi'') = \int g_r(\xi'\xi''')\, d\xi'''\, .\, (\xi'''/g')\,(g'/\xi'') = g_r(\xi' g') \cdot (g'/\xi'')$$

$$= g_r' \cdot (\xi'/g') \cdot (g'/\xi'') = g_r'\, \delta(g - g')\,(\xi'\xi''),$$

and similarly

$$\delta(g - g')\, g_r(\xi'\xi'') = g_r'\, \delta(g - g')\,(\xi'\xi'').$$

The volume of η-space for which the g's lie between their given numerical values when $\xi_r = \xi_r'$ is now given by

$$\int_{g'}^{g''} (\xi'/g')\, dg'\, (g'/\xi').$$

We see at once that this volume vanishes unless the range of integration of each g_r includes a characteristic value of that g_r [*i.e.*, a value that occurs as a diagonal element of the matrix representing g_r in the (g) scheme of matrix

* Where g_r' is used in a formula as a matrix it means the diagonal matrix, with elements $g_r'\,(\xi'\xi'') = g_r'\,\delta(\xi' - \xi'')$, which represents the c-number g_r'.

representation], since otherwise the transformation functions (ξ'/g') and (g'/ξ') would vanish throughout the range of integration. This shows that the characteristic values of any constant of integration g are the values (quantised or otherwise) that this q-number can actually take. (In particular, the characteristic values of H are the energy levels of the system.) The symmetry between ξ' and g' in the (ξ') diagonal element, namely $(\xi'/g')\,(g'/\xi')$, of the matrix $\delta(g-g')$ enables one to formulate a reciprocal theorem of quantum dynamics, applying only to the case when the transformation functions (ξ'/g'), (g'/ξ') are continuous functions of the ξ_r' and g_r' (which implies that the characteristic values of the ξ_r and g_r can take continuous ranges of values) as follows :— The volume of η-space for which $\xi_r = \xi_r'$ and $g_r' < g_r < g_r' + \varepsilon_r$ is equal to the volume of the space of the variables canonically conjugate to the g's for which $g_r = g_r'$ and $\xi_r' < \xi_r < \xi_r' + \varepsilon_r$, where the ε_r are small positive c-numbers. Either of these volumes is, in fact, just $(\xi'/g')\,(g'/\xi')\,\varepsilon$, where ε is the product of all the ε_r.

§ 7. *Comparison with Previous Methods.*

We shall now show that the present method of obtaining physical results from the matrix theory is in agreement with the assumptions formerly used that the square of the amplitude of the wave function in certain cases determines a probability. Consider a dynamical system which, when unperturbed, has a Hamiltonian which does not involve the time explicitly, and to which a perturbation is applied, causing an additional term, that does involve the time explicitly, to appear in the Hamiltonian. To find the transition probabilities induced by the perturbation, according to the former method, one must first obtain the eigenfunctions $\psi_0(\alpha')$, say, for the unperturbed system, (the α's being constants of integration of the unperturbed system), and then the eigenfunctions $\psi_t(\alpha')$, say, that satisfy the wave equation of the perturbed system and have the initial values $\psi_0(\alpha')$. One must then expand the ψ_t's in terms of the ψ_0's thus,

$$\psi_t(\alpha') = \int \psi_0(\alpha'')\, d\alpha''\, c(\alpha''\alpha'), \tag{15}$$

where the coefficients $c(\alpha''\alpha')$ are functions of the time only. One then assumes that $|c(\alpha''\alpha')|^2\, d\alpha''$ is the probability of an atom initially in the state (α') being at the time t in a state for which each α_r lies between α_r'' and $\alpha_r'' + d\alpha_r''$.

To determine this probability by means of the general method of the present paper, we must find the transformation functions (α_t'/α_0') and (α_0'/α_t') which connect the values α_t of the variables α (which are assumed to be functions of the p's and q's that do not involve the time explicitly) at the time t with their

initial values α_0, both the α_t's and the α_0's being constants of integration of the perturbed system if we regard t as a specified time. The required probability would then be equal to

$$(\alpha_0'/\alpha_t')\, d\alpha_t'\, (\alpha_t'/\alpha_0') = |(\alpha_t'/\alpha_0')|^2\, d\alpha_t'$$

if (α_0'/α_t') and (α_t'/α_0') are conjugate imaginaries, as they have to be if both the (α_t) and the (α_0) schemes of matrix representation satisfy the condition (iv) of § 3. If by q_t we denote the value of each co-ordinate q at the specified time t, we have

$$(q_t'/\alpha_0') = \int (q_t'/\alpha_t')\, d\alpha_t'\, (\alpha_t'/\alpha_0'). \tag{15'}$$

Here (q_t'/α_0') is the eigenfunction that satisfies the perturbed wave equation [of the form (12)], and (q_t'/α_t'), which depends only on the analytical relations connecting the α's with the p's and q's, is the same function of the q_t''s and α_t''s that (q_0'/α_0') is of the q_0''s (the initial q's) and α_0''s, and is therefore the eigenfunction for the undisturbed system written in terms of the variables q_t', α_t'. Equation (15') is thus the same as equation (15), and the transformation functions (α_t'/α_0') are just the coefficients of equation (15) [which ought really to have been written $c(\alpha_t''\alpha_0')$]. Hence the present general method gives results identical with those of the previous assumption.

Consider now the case of encounters between, say, an electron and an atomic system. In Born's treatment of the problem one finds a solution of Schrödinger's wave equation consisting of incident plane waves representing the approaching electron, which waves are scattered by the atomic system. One then assumes that the square of the amplitude of the wave scattered in any direction determines the probability of the electron being scattered in that direction, with an energy given by the frequency of the wave.

To determine this probability by the present method, we must find the transformation function $(p_{\mathrm{F}}'/p_{\mathrm{I}}')$ that connects the final components of momentum of the electron, p_{F}, with the initial components p_{I}. There is then a probability $(p_{\mathrm{I}}'/p_{\mathrm{F}}')\, dp_{\mathrm{F}}'\, (p_{\mathrm{F}}'/p_{\mathrm{I}}') = |(p_{\mathrm{F}}'/p_{\mathrm{I}}')|^2\, dp_{\mathrm{F}}'$ that the electron will be scattered with a momentum lying in the range dp_{F}'. If the co-ordinates of the electron at the time t are x_t,* we have

$$(x_t'/p_{\mathrm{I}}') = \int (x_t'/p_{\mathrm{F}}')\, dp_{\mathrm{F}}'\, (p_{\mathrm{F}}'/p_{\mathrm{I}}'). \tag{16}$$

* The set of q-numbers x_t is understood, of course, to include the co-ordinates of the atomic system, which are not explicitly mentioned, in addition to the co-ordinates x_t of the incident electron. In the same way the p_{I} and p_{F} must include variables to fix the stationary states of the atomic system.

Here the transformation function (x_t'/p_I') is the solution of Schrödinger's wave equation appropriate for the case of an incident electron with momentum p_I', and is thus the wave function of Born's theory. The function (x_t'/p_F') on the other hand represents emerging waves corresponding to electrons with momentum p_F' (and also in-going waves that we need not consider). Equation (16) thus gives the resolution of the emerging waves in the eigenfunction (x_t'/p_I') into their different components, the amplitudes of the various components being $|(p_F'/p_I')|$. The present method is therefore in agreement with Born's theory.

If these problems are regarded from the matrix point of view, one sees that the dynamical variables must be capable of being represented equally well by matrices whose rows and columns refer to the initial values of the action variables (α_0 or p_I in the two cases) or to the final values (α_t or p_F), and the coefficients that enable one to transform from the one set of matrices to the other are just those that determine the transition probabilities.

In conclusion it may be mentioned that the present theory suggests a point of view for regarding quantum phenomena rather different from the usual ones. One can suppose that the initial state of a system determines definitely the state of the system at any subsequent time. If, however, one describes the state of the system at an arbitrary time by giving numerical values to the co-ordinates and momenta, then one cannot actually set up a one-one correspondence between the values of these co-ordinates and momenta initially and their values at a subsequent time. All the same one can obtain a good deal of information (of the nature of averages) about the values at the subsequent time considered as functions of the initial values. The notion of probabilities does not enter into the ultimate description of mechanical processes; only when one is given some information that involves a probability (*e.g.*, that all points in η-space are equally probable for representing the system) can one deduce results that involve probabilities.