On the Relation between the Quantum Mechanics of Heisenberg, Born, and Jordan, and that of Schrödinger

(*Annalen der Physik* (4), vol. 79, 1926)

§ 1. Introduction and Abstract

CONSIDERING the extraordinary differences between the starting-points and the concepts of Heisenberg's quantum mechanics [1] and of the theory which has been designated "undulatory" or "physical" mechanics,[2] and has lately been described here, it is very strange that these two new theories agree *with one another* with regard to the known facts, where they differ from the old quantum theory. I refer, in particular, to the peculiar "half-integralness" which arises in connection with the oscillator and the rotator. That is really very remarkable, because starting-points, presentations, methods, and in fact the whole mathematical apparatus, seem fundamentally different. Above all, however, the departure from classical mechanics in the two theories seems to occur in diametrically opposed directions. In Heisenberg's work the classical continuous variables are replaced by systems of discrete numerical quantities (matrices), which depend on a pair of integral indices, and are defined by *algebraic* equations. The authors themselves describe the theory as a "true theory of a discontinuum".[3] On the other hand, wave mechanics shows just the reverse tendency; it is a step from classical point-mechanics towards a *continuum-theory*. In place of a process described in terms of a finite number of dependent variables occurring in a finite number of total differential equations, we have a continuous *field-like* process in

[1] W. Heisenberg, *Ztschr. f. Phys.* 33, p. 879, 1925; M. Born and P. Jordan, *idem* 34, p. 858, 1925, and 35, p. 557, 1926 (the latter in collaboration with Heisenberg). I may be allowed, for brevity's sake, to replace the three names simply by Heisenberg, and to quote the last two essays as "Quantum Mechanics I. and II." Interesting contributions to the theory have also been made by P. Dirac, *Proc. Roy. Soc.*, London, 109, p. 642, 1925, and *idem* 110, p. 561, 1926.

[2] E. Schrödinger. Parts I. and II. in this collection. These parts will be continued quite independently of the present paper, which is only intended to serve as a connecting link.

[3] "Quantum Mechanics I." p. 879.

configuration space, which is governed by a single *partial* differential equation, derived from a principle of action. This principle and this differential equation replace the equations of motion *and* the quantum conditions of the older " classical quantum theory ".[1]

In what follows the very intimate *inner connection* between Heisenberg's quantum mechanics and my wave mechanics will be disclosed. From the formal mathematical standpoint, one might well speak of the *identity* of the two theories. The train of thought in the proof is as follows.

Heisenberg's theory connects the solution of a problem in quantum mechanics with the solution of a system of an infinite number of algebraic equations, in which the unknowns—infinite matrices—are allied to the classical position- and momentum-co-ordinates of the mechanical system, and functions of these, and obey peculiar *calculating rules*. (The relation is this : to *one* position-, *one* momentum-co-ordinate, or to *one* function of these corresponds always *one* infinite matrix.)

I will first show (§§ 2 and 3) how to each function of the position- and momentum-co-ordinates there may be related a matrix in such a manner, that these matrices, *in every case, satisfy* the formal calculating rules of Born and Heisenberg (among which I also reckon the so-called " quantum condition " or " interchange rule " ; see below). This relation of matrices to functions is *general* ; it takes no account of the *special* mechanical system considered, but is the same for all mechanical systems. (In other words : the particular Hamilton function does not enter into the connecting law.) However, the relation is still *indefinite* to a great extent. It arises, namely, from the *auxiliary introduction* of an *arbitrary* complete orthogonal system of functions having for domain *entire configuration space* (*N.B.—not* " *pq*-space ", but " *q*-space "). The provisional *indefiniteness* of the relation lies in the fact that we can assign the *auxiliary* rôle to an *arbitrary* orthogonal system.

After matrices are thus constructed in a very general way, so as to satisfy the general rules, I will show the following in § 4. The *special* system of algebraic equations, which, in a *special* case, connects the *matrices* of the position and impulse co-ordinates with the *matrix* of the Hamilton function, and which the authors call " equations of motion ", will be completely solved by assigning the auxiliary rôle to a *definite* orthogonal system, namely, to the system of *proper functions* of that partial differential equation which forms the basis of my wave mechanics. The solution of the natural *boundary-value problem* of this differential equation is *completely equivalent* to the solution of Heisenberg's algebraic problem. *All* Heisenberg's matrix elements, which

[1] My theory was inspired by L. de Broglie, *Ann. de Physique* (10) 3, p. 22, 1925 (*Theses*, Paris, 1924), and by brief, yet infinitely far-seeing remarks of A. Einstein, *Berl. Ber.*, 1925, p. 9 *et seq.* I did not at all suspect any relation to Heisenberg's theory at the beginning. I naturally knew about his theory, but was discouraged, if not repelled, by what appeared to me as very difficult methods of transcendental algebra, and by the want of perspicuity (*Anschaulichkeit*).

may interest us from the surmise that they define "transition probabilities" or "line intensities", can be actually evaluated *by differentiation and quadrature*, as soon as the *boundary-value problem* is solved. Moreover, in wave mechanics, these matrix elements, or quantities that are closely related to them, have the perfectly clear significance of amplitudes of the partial oscillations of the atom's electric moment. The intensity and polarisation of the emitted light is thus intelligible *on the basis of the Maxwell-Lorentz theory*. A short preliminary sketch of this relationship is given in § 5.

§ 2. The Co-ordination of an Operator and of a Matrix with a Well-arranged Function-symbol and the Establishment of the Product Rule

The starting-point in the construction of matrices is given by the simple observation that Heisenberg's peculiar calculating laws for functions of the *double* set of n quantities, $q_1, q_2, \ldots, q_n$; $p_1, p_2, \ldots, p_n$ (position- and canonically conjugate momentum-co-ordinates) agree exactly with the rules, which *ordinary analysis* makes *linear differential operators* obey in the domain of the *single* set of n variables, $q_1, q_2, \ldots, q_n$. So the *co-ordination* has to occur in such a manner that each p_l in the *function* is to be replaced by the operator $\frac{\partial}{\partial q_l}$. Actually the operator $\frac{\partial}{\partial q_l}$ is *exchangeable* with $\frac{\partial}{\partial q_m}$, where m is arbitrary, but with q_m only, if $m \neq l$. The operator, obtained by interchange and subtraction when $m = l$, viz.

$$\frac{\partial}{\partial q_l} q_l - q_l \frac{\partial}{\partial q_l}, \tag{1}$$

when applied to any arbitrary function of the q's, *reproduces* the function, *i.e.* this operator gives *identity*. This simple fact will be reflected in the domain of matrices as Heisenberg's interchange rule.

After this preliminary survey, we turn to systematic construction. Since, as noticed above, the interchangeability does not *always* hold good, then a definite operator does not correspond uniquely to a definite "function in the usual sense" of the q's and p's, but to a "function-symbol written in a definite way". Moreover, since we can perform only the operations of addition and multiplication with the operators $\frac{\partial}{\partial q_k}$, the function of the q's and p's must be written as a regular power series in p at least, before we substitute $\frac{\partial}{\partial q_l}$ for p_l. It is sufficient to carry out the process for a single term of such a power series, and thus for a function of the following construction:

$$F(q_k, p_k) = f(q_1 \ldots q_n) p_r p_s p_t g(q_1 \ldots q_n) p_{r'} \; h(q_1 \ldots q_n) p_{r'} p_{s'} \ldots \tag{2}$$

We wish to express this as a "well-arranged[1] function-symbol" and relate it to the following operator,

[1] Or "well-ordered."

$$(3)\quad [F,\ \cdot\ ]=f(q_1 \ldots q_n)K^3\frac{\partial^3}{\partial q_r \partial q_s \partial q_i}g(q_1 \ldots q_n)K\frac{\partial}{\partial q_r}$$
$$h(q_1 \ldots q_n)K^2\frac{\partial^2}{\partial q_r \partial q_s}\ldots,$$

wherein, somewhat more generally than in the preliminary survey, p_r is not replaced by $\frac{\partial}{\partial q_r}$ simply, but by $K\frac{\partial}{\partial q_r}$, and K stands for a universal constant. As an abbreviation for the operator arising out of the well-arranged function F, I have introduced the symbol $[F,\ \cdot\ ]$ in passing (*i.e.* only for the purpose of the present proof). The function (in the usual sense) of $q_1 \ldots q_n$, which is obtained by using the operator on another function (in the usual sense), $u(q_1 \ldots q_n)$, will be denoted by $[F, u]$. If G is another well-arranged function, then $[GF, u]$ will denote the function u after the operator of F has *first* been used on it, and *then* the operator of G; or, what is defined to be the same, when the operator of GF has been used. Of course this is not generally the same as $[FG, u]$.

Now we connect a *matrix* with a well-arranged function, like F, by means of its operator (3) and of an arbitrary complete orthogonal system having for its domain the whole of q-space. It is done as follows. For brevity we will simply write x for the group of variables $q_1, q_2, \ldots q_n$, as is usual in the theory of Integral Equations, and write $\int dx$ for an integral extending over the whole of q-space. The functions

$$(4)\qquad u_1(x)\sqrt{\rho(x)},\quad u_2(x)\sqrt{\rho(x)},\quad u_3(x)\sqrt{\rho(x)}\ \ldots\ \text{ad inf.}$$

are now to form a complete orthogonal system, normalised to 1.

Let, therefore, in every case

$$(5)\qquad \begin{cases} \int \rho(x)u_i(x)u_k(x)dx = 0 \text{ for } i \neq k \\ \qquad\qquad\qquad\quad = 1 \text{ for } i = k. \end{cases}$$

Further, it is postulated that these functions vanish at the natural *boundary* of q-space (in general, infinity) in a way sufficient to cause the vanishing of certain boundary integrals which come in later on as secondary products after certain integrations by parts.

By the operator (3) we now relate the following *matrix*,

$$(6)\qquad F^{kl} = \int \rho(x)u_k(x)[F, u_l(x)]dx,$$

to the function F represented by (2). (The way of writing the indices on the left-hand side must not suggest the idea of "contravariance"; from this point of view, here discarded, *one* index was formerly written above, and the other below; we write the matrix indices *above*, because later we will also have to write matrix elements, corresponding to the q's and p's, where the lower place is already occupied.) In words: a matrix element is computed by *multiplying* the function of

the orthogonal system denoted by the *row*-index (whereby we understand always u_i, not $u_i\sqrt{\rho}$), by the "density function" ρ, and by the result arising from using our operator on the orthogonal function corresponding to the *column*-index, and then by *integrating* the whole over the domain.[1]

It is not very difficult to show that additive and multiplicative combination of well-arranged functions or of the appertaining operators works out as matrix addition and matrix multiplication of the allied matrices. For addition the proof is trivial. For multiplication the proof runs as follows. Let G be any other well-arranged function, like F, and

$$G^{lm} = \int \rho(x) u_l(x) [G, u_m(x)] dx, \tag{7}$$

the matrix corresponding. We wish to form the product matrix

$$(FG)^{km} = \sum_l F^{kl} G^{lm}.$$

Before writing it, let us transform the expression (6) for F^{kl} as follows. By a series of integrations by parts, the operator $[F, \cdot\,]$ is "revolved" from the function $u_l(x)$ to the function $\rho(x)u_k(x)$. By the expression "revolve" (instead of, say, "push") I wish to convey that the *sequence* of the operations reverses itself exactly thereby. The boundary integrals, which come in as "by-products", are to disappear (see above). The "revolved" operator, including the change of sign that accompanies an odd number of differentiations, will be denoted by $[\overline{F}, \cdot\,]$. For example, from (3) comes

$$[\overline{F}, \cdot\,] = (-1)^\tau \ldots K^2 \frac{\partial^2}{\partial q_s \partial q_{s'}} h(q_1 \ldots q_n) K \frac{\partial}{\partial q_{s'}} \tag{3'}$$
$$g(q_1 \ldots q_n) K^3 \frac{\partial^3}{\partial q_t \partial q_s \partial q_r} f(q_1 \ldots q_n),$$

where τ = number of differentiations. By applying this symbol, we have

$$F^{kl} = \int u_l(x) [\overline{F}, \rho(x) u_k(x)] dx. \tag{6'}$$

If we now calculate the product matrix, we get

$$\sum_l F^{kl} G^{lm} \tag{8}$$
$$= \sum_l \left\{ \int u_l(x) [\overline{F}, \rho(x) u_k(x)] dx \cdot \int \rho(x) u_l(x) [G, u_m(x)] dx \right\}$$
$$= \int [\overline{F}, \rho(x) u_k(x)] [G, u_m(x)] dx.$$

The last equation is simply the so-called "relation of completeness"

[1] More briefly: F^{kl} is the kth "development coefficient" of the operator used on the function u_l.

of our orthogonal system,[1] applied to the "development coefficients" of the functions

$$[G, u_m(x)] \quad \text{and} \quad \frac{1}{\rho(x)}[F, \rho(x)u_k(x)].$$

Now in (8), let us revolve, by further integrations by parts, the operator $[\overline{F}, \cdot\,]$ from the function $\rho(x)u_k(x)$ back again to the function $[G, u_m(x)]$, so that the operator regains its original form. We clearly get

$$(FG)^{km} = \sum_l F^{kl}G^{lm} = \int \rho(x)u_k(x)[FG, u_m(x)]dx. \tag{9}$$

On the left is the (km)th element of the product matrix, and on the right, by the law of connection (6), stands the (km)th element of the matrix, corresponding to the well-arranged product FG. Q.E.D.

§ 3. Heisenberg's Quantum Condition and the Rules for Partial Differentiation

Since operation (1) gave identity, then corresponding to the well-arranged function

$$p_l q_l - q_l p_l \tag{10}$$

we have the *operator*, multiplication by K, in accordance with our law of connection, in which we incorporated a universal constant K. Hence to function (10) corresponds the *matrix*

$$(p_l q_l - q_l p_l)^{ik} = K\int \rho(x)u_i(x)u_k(x)dx = 0 \text{ for } i \neq k \tag{11}$$
$$= K \text{ for } i = k.$$

That is Heisenberg's "quantum relation" if we put

$$K = \frac{h}{2\pi\sqrt{-1}}, \tag{12}$$

and this may be assumed to hold from now on. It is understood that we could have also found relation (11) by taking the two matrices allied to q_l and p_l, viz.

$$q_l^{ik} = \int q_l \rho(x)u_i(x)u_k(x)dx, \tag{13}$$
$$p_l^{ik} = K\int \rho(x)u_i(x)\frac{\partial u_k(x)}{\partial q_l}dx,$$

multiplying them together in different sequence and subtracting the two results.

Let us now turn to the "rules for partial differentiation". A well-arranged function, like (2), is said to be differentiated partially with respect to q_l, when it is differentiated with respect to q_l without

[1] See, *e.g.*, Courant-Hilbert, *Methods of Mathematical Physics*, I., p. 36. It is important to remember that the "relation of completeness" for the "development coefficients" is valid in every case, even when the developments themselves do *not* converge. If these do converge, then the equivalence (8) is directly evident.

altering the succession of the factors *at each place* where q_l appears in it, and all these results are added.[1] Then it is easy to show that the following equation between the operators is valid:

$$\left[\frac{\partial F}{\partial q_l}, \cdot\right] = \frac{1}{K}[p_l F - F p_l, \cdot]. \tag{14}$$

The line of thought is this. Instead of really differentiating with respect to q_l, it is very convenient simply to prefix p_l to the function; as it is, p_l must finally be replaced by $K\frac{\partial}{\partial q_l}$. Obviously I have to divide by K. Furthermore, when we apply the entire operator to any function u, the operator $\frac{\partial}{\partial q_l}$ will act not only on that part of F which contains q_l (as it *ought*), but also *wrongly* on the function u, affected by the entire operator. *This mistake is exactly corrected* by subtracting again the operation $[F p_l, \cdot]$!

Consider now partial differentiation with respect to a p_l. Its meaning for a well-arranged function, like (2), is a little simpler than in the case of $\frac{\partial}{\partial q_l}$, because the p's only appear as power products. We imagine every power of p_l to be resolved into single factors, *e.g.* think of $p_l p_l p_l$ instead of p_l^3, and we can then say: in partial differentiation with respect to p_l, every *separate* p_l that appears in F is to be *dropped* once, all the other p_l's remaining; all the results obtained are to be added. What will be the effect on the operator (3)? "Every separate $K\frac{\partial}{\partial q_l}$ is to be dropped once, and all the results so obtained are to be added."

I maintain that on this reasoning the operational equation

$$\left[\frac{\partial F}{\partial p_l}, \cdot\right] = \frac{1}{K}[F q_l - q_l F, \cdot] \tag{15}$$

is valid. Actually, I picture the operator $[F q_l, \cdot]$ as formed and now attempt to "push q_l through F from right to left", that means, attempt to arrive at the operator $[q_l F, \cdot]$ through successive exchanges. This pushing through meets an obstacle only as often as I come against a $\frac{\partial}{\partial q_l}$. With the latter I may not interchange q_l simply, but have to replace

$$\frac{\partial}{\partial q_l} \quad \text{by} \quad 1 + q_l \frac{\partial}{\partial q_l} \tag{16}$$

in the interior of the operator. The secondary products of the interchange, which are yielded by this "uniformising", form just the

[1] We are naturally following Heisenberg faithfully in all these definitions. From a strictly logical standpoint the following proof is evidently superfluous, and we could have written down rules (14) and (15) right away, as they are proved in Heisenberg, and only depend upon the sum and product rules and the exchange rule (11) which we have proved.

desired " partial differential coefficients ", as is easily seen. After the pushing-through process is finished, the operator $[q_l F, \cdot\,]$ still remains left over. It would be superfluous and therefore is explicitly subtracted in (15). Hence (15) is proved. The equations (14) and (15), which have been proved for *operators*, naturally hold good unchanged for the matrices belonging to the right-hand and left-hand sides, because by (6) one matrix, and one only, belongs to one linear operator (after the system $u_i(x)$ has been chosen once for all).[1]

§ 4. The Solution of Heisenberg's Equations of Motion

We have now shown that matrices, constructed according to definitions (3) and (6) from well-arranged functions by the agency of an arbitrary, complete orthogonal system (4), satisfy all Heisenberg's calculating rules, including the interchange rule (11). Now let us consider a special mechanical problem, characterised by a definite Hamilton function

$$H(q_k, p_k). \tag{17}$$

The authors of quantum mechanics take this function over from *ordinary* mechanics, which naturally does *not* give it in a " well-arranged " form; for in ordinary analysis no stress is laid on the sequence of the factors. They therefore " normalise " or " symmetricalise " the function in a definite manner for their purposes. For example, the usual mechanical function $q_k p_k^2$ is replaced by

$$\tfrac{1}{2}(p_k^2 q_k + q_k p_k^2)$$

[1] In passing it may be noted that the converse of this theorem is also true, at least in the sense that certainly *not more* than *one* linear differential operator can belong to a given *matrix*, according to our connecting law (6), when the orthogonal system and the density function are prescribed. For in (6), let the F^{kl}'s be given, let $[F, \cdot\,]$ be the linear operator we are *seeking* and which we *presume* to *exist*, and let $\phi(x)$ be a function of $q_1, q_2, \ldots, q_n$, which is sectionally continuous and differentiable as often as necessary, but otherwise arbitrary. Then the *relation of completeness* applied to the functions $\phi(x)$ and $[F, u_k(x)]$ yields the following:

$$\int \rho(x)\phi(x)[F, u_k(x)]dx = \sum_l \left\{ \int \rho(x)\phi(x)u_l(x)dx \cdot \int \rho(x)u_l(x)[F, u_k(x)]dx \right\}.$$

The right-hand side can be regarded as definitely known, for in it occur only development coefficients of $\phi(x)$ and the prescribed matrix elements F^{lk}. By " revolving " (see above), we can change the left-hand side into the kth development coefficient of the function

$$\frac{[\overline{F}, \rho(x)\phi(x)]}{\rho(x)}.$$

Thus all the development coefficients of this function are uniquely fixed, and thus so is the function itself (Courant-Hilbert, p. 37). Since, however, $\rho(x)$ was fixed beforehand and $\phi(x)$ is a quite arbitrary function, we can say: the result of the action of the *revolved* operator on an *arbitrary* function, provided, of course, it can be submitted to the operator at all, is fixed *uniquely* by the matrix F^{kl}. This can only mean that *the revolved operator* is uniquely fixed, for the notion of " operator " is logically identical with the whole of the results of its action. By revolving the revolved operator, we obtain uniquely the operator we have sought, itself.

It is to be noted that the *developability* of the functions which appear is *not* necessarily postulated—we have not proved that a linear operator, corresponding to an arbitrary matrix, *always exists*.

or by $$p_k q_k p_k$$

or by $$\tfrac{1}{3}(p_k^2 q_k + p_k q_k p_k + q_k p_k^2),$$

which are all the same, according to (11). This function is then " well-arranged ", *i.e.* the sequence of the factors is inviolable. I will not enter into the general rule for symmetricalising here ; [1] the idea, if I understand it aright, is that H^{ki} is to be a *diagonal matrix*, and in other respects the normalised function, regarded as one of ordinary analysis, is to be identical with the one originally given.[2] We will satisfy these demands in a direct manner.

Then the authors postulate that the *matrices* q_l^{ik}, p_l^{ik} shall satisfy an infinite system of equations, as " equations of motion ", and to begin with they write this system as follows :

$$(18)\qquad \left.\begin{aligned} \left(\frac{dq_l}{dt}\right)^{ik} &= \left(\frac{\partial H}{\partial p_l}\right)^{ik} \\ \left(\frac{dp_l}{dt}\right)^{ik} &= \left(-\frac{\partial H}{\partial q_l}\right)^{ik} \end{aligned}\right\} \begin{aligned} &l = 1, 2, 3, \ldots n \\ &i, k = 1, 2, 3, \ldots \text{ad. inf.} \end{aligned}$$

The upper pair of indices signifies, as before in F^{ki}, the respective element of the matrix belonging to the well-arranged function in question. The meaning of the partial differential coefficient on the right-hand side has just been explained, but *not* that of the $\frac{d}{dt}$ appearing on the left. By it the authors signify the following. *It* is to *give* a series of *numbers*

$$(19)\qquad \nu_1, \nu_2, \nu_3, \nu_4, \ldots \text{ ad inf.},$$

such that the above equations are fulfilled, when to the $\frac{d}{dt}$ is ascribed the meaning : multiplication of the (ik)th matrix element by $2\pi\sqrt{-1}(\nu_i - \nu_k)$. Thus, in particular,

$$(20)\qquad \begin{cases} \left(\frac{dq_l}{dt}\right)^{ik} = 2\pi\sqrt{-1}(\nu_i - \nu_k) q_l^{ik}\,; \\ \left(\frac{dp_l}{dt}\right)^{ik} = 2\pi\sqrt{-1}(\nu_i - \nu_k) p_l^{ik}. \end{cases}$$

The series of numbers (19) is not defined in any way beforehand, but together with the matrix elements q_l^{ik}, p_l^{ik}, they form the numerical unknowns of the system of equations (18). The latter assumes the form

$$(18')\qquad \begin{cases} (\nu_i - \nu_k) q_l^{ik} = \frac{1}{h}(Hq_l - q_l H) \\ (\nu_i - \nu_k) p_l^{ik} = \frac{1}{h}(Hp_l - p_l H) \end{cases}$$

[1] " Quantum Mechanics I." p. 873 *et seq.*

[2] The *stricter* postulation—"shall yield the same quantum-mechanical equations of motion "—I consider too narrow. It arises, in my opinion, from the fact that the authors confine themselves to *power products with regard also to the* q_k's—which is unnecessary.

when we utilise the explanation of the symbols (20), and the calculating rules (14) and (15), and take account of (12).

We must thus satisfy *this* system of equations, and we have no means at our disposal, other than the suitable choice of the orthogonal system (4), which intervenes in the formation of the matrices. I now assert the following:

1. The equations (18′) will in general be satisfied if we choose as the orthogonal system the *proper functions* of the natural boundary value problem of the following partial differential equation,

$$-[H, \psi] + E\psi = 0. \tag{21}$$

ψ is the unknown function of $q_1, q_2, \ldots, q_n$; E is the proper value parameter. Of course, as density function, $\rho(x)$ appears that function of $q_1, \ldots, q_n$, by which equation (21) must be multiplied in order to make it self-adjoint. The quantities ν_i are found to be equal to the proper values E_i divided by h. H^{kl} becomes a diagonal matrix, with $H^{kk} = E_k$.

2. If the symmetricalising of the function H has been effected *in a suitable way*—the process of symmetricalising, in my opinion, has not hitherto been defined uniquely—then (21) *is identical with the wave equation which is the basis of my wave mechanics.*[1]

Assertion 1 is almost directly evident, if we provisionally lay aside the questions whether equation (21) gives rise at all to an intelligible boundary value problem with the domain of entire q-space, and whether it can always be made self-adjoint through multiplication by a suitable function, etc. These questions are largely settled under heading 2. For now we have, according to (21) and the definitions of proper values and functions,

$$[H, u_i] = E_i u_i, \tag{22}$$

and thus from (6) we get

$$\begin{cases} H^{kl} = \int \rho(x) u_k(x) [H, u_l(x)] dx = E_i \int \rho(x) u_k(x) u_l(x) dx \\ \quad = 0 \text{ for } l \neq k \\ \quad = E_l \text{ for } l = k, \end{cases} \tag{23}$$

and, for example,

$$\begin{cases} (Hq_l)^{ik} = \sum_m H^{im} q_l^{mk} = E_i q_l^{ik} \\ (q_l H)^{ik} = \sum_m q_l^{im} H^{mk} = E_k q_l^{ik}, \end{cases} \tag{24}$$

so that the right-hand side of the first equation of (18′) takes the value

$$\frac{E_i - E_k}{h} q_l^{ik}. \tag{25}$$

Similarly for the second equation. Thus everything asserted under 1 is proved.

[1] Equation (18″), Part II.

Let us turn now to assertion 2, which is, that there is agreement between the negatively taken operator of the Hamilton function (suitably symmetricalised) and the wave operator of wave mechanics. I will first illustrate by a simple example why the process of symmetricalisation seems to me to be, *in the first instance*, not unique. Let, for *one* degree of freedom, the *ordinary* Hamilton function be

$$H = \tfrac{1}{2}(p^2 + q^2). \tag{26}$$

Then it is admitted that we can take this function, just as it stands, unchanged, over to "quantum mechanics" as a "well-arranged" function. But we can also, and seemingly indeed with as much right *to begin with*, apply the well-arranged function

$$H = \tfrac{1}{2}\left(\frac{1}{f(q)}pf(q)p + q^2\right), \tag{27}$$

where $f(q)$ is a function arbitrary within wide limits. $f(q)$ would appear in this case as a "density function" $\rho(x)$. (26) is quite evidently just a special case of (27), and the question arises, whether (and how) it is at all possible to distinguish the special case we are concerned with, *i.e.* for more complicated H-functions. Confining ourselves to power products only of the q_k's (where we could then simply prohibit the "production of denominators") would be most inconvenient just in the most important applications. Besides, I believe that does *not* lead to correct symmetricalisation.

For the convenience of the reader, I will now give again a short derivation of the wave equation in a form suited to the present purpose, confining myself to the case of classical mechanics (without relativity and magnetic fields). Let, therefore,

$$H = T(q_k, p_k) + V(q_k), \tag{28}$$

T being a quadratic form in the p_k's. Then the wave equation can be deduced [1] from the following variation problem,

$$\begin{cases} \delta J_1 = \delta \int \left\{ \dfrac{h^2}{4\pi^2} T\left(q_k, \dfrac{\partial \psi}{\partial q_k}\right) + \psi^2 V(q_k) \right\} \Delta_p^{-\frac{1}{2}} dx = 0, \\ \text{with the subsidiary condition} \\ J_2 = \int \psi^2 \Delta_p^{-\frac{1}{2}} dx = 1. \end{cases} \tag{29}$$

As above, $\int dx$ stands for $\int \ldots \int dq_1 \ldots dq_n$; $\Delta_p^{-\frac{1}{2}}$ is *the reciprocal of the square root of the discriminant* of the quadratic form T. *This factor must not be omitted*, because otherwise the whole process would not be invariant for point transformations of the q's! By all means another *explicit* function of the q's might appear as a factor, *i.e.* a function which would be invariant for a point transformation of the q's. (For Δ_p, as is known, this is not the case. Otherwise we *could* omit $\Delta_p^{-\frac{1}{2}}$, if this extra function was given the value $\Delta_p^{\frac{1}{2}}$.)

If we indicate the derivative of T with respect to *that* argument,

[1] Equations (23) and (24) of Part I.

which originally was p_k, by the suffix p_k, we obtain, as the *result of the variation*,

$$(30)\qquad \left\{\begin{aligned} 0 &= \tfrac{1}{2}(\delta J_1 - E\delta J_2) \\ &= \int\left\{-\frac{h^2}{8\pi^2}\sum_k \frac{\partial}{\partial q_k}\left[\Delta_p^{-\frac{1}{2}} T_{p_k}\left(q_k, \frac{\partial\psi}{\partial q_k}\right)\right]\right. \\ &\qquad \left. + \left(V(q_k) - E\right)\Delta_p^{-\frac{1}{2}}\psi\right\}\delta\psi dx; \end{aligned}\right.$$

the Eulerian variation equation thus runs:

$$(31)\qquad \frac{h^2}{8\pi^2}\Delta_p^{\frac{1}{2}}\sum_k \frac{\partial}{\partial q_k}\left\{\Delta_p^{-\frac{1}{2}} T_{p_k}\left(q_k, \frac{\partial\psi}{\partial q_k}\right)\right\} - V(q_k)\psi + E\psi = 0.$$

It is not difficult to see that this equation has the form of (21) if we remember our law connecting the operators, and consider

$$(32)\qquad T(q_k, p_k) = \tfrac{1}{2}\sum_k p_k T_{p_k}(q_k, p_k)$$

the Eulerian equation for homogeneous functions, applied to the quadratic form T. In actual fact, if we detach the operator from the left side of (31), with the proper value term $E\psi$ removed, and replace in it $\dfrac{h}{2\pi\sqrt{-1}}\dfrac{\partial}{\partial q_k}$ by p_k, then according to (32) we obtain the negatively taken Hamilton function (28). Thus the process of variation has given quite automatically a uniquely defined "symmetricalisation" of the operator, which makes it self-adjoint (except possibly for a common factor) and makes it invariant for point transformations, and which I would like to maintain, as long as there are no definite reasons for the appearance under the integrals (29) of the additional factor, already [1] mentioned as possible, and for a definite form of the latter.

Hence the solution of the whole system of matrix equations of Heisenberg, Born, and Jordan is reduced to the natural boundary value problem of a linear partial differential equation. If we have solved the boundary value problem, then by the use of (6) we can calculate by differentiations and quadratures every matrix element we are interested in.

As an illustration of what is to be understood by the *natural* boundary value problem, *i.e.* by the natural boundary conditions at the natural boundary of configuration space, we may refer to the worked examples.[2] It invariably turns out that the natural infinitely distant boundary forms a singularity of the differential equation and only allows of the one boundary condition—"remaining finite". This seems to be a general characteristic of those micro-mechanical problems with which the theory in the first place is meant to deal. If the domain of the position co-ordinates is artificially limited (example: a molecule in a "vessel"), then an essential allowance must be made for this limitation by the introduction of suitable potential energies in

[1] Cf. also *Ann. d. Phys.* 79, p. 362 and p. 510 (*i.e.* Parts I. and II.).
[2] In Parts I. and II. of this collection.

the well-known manner. Also the *vanishing* of the proper functions at the boundary generally occurs to an adequate degree, even if relations among *certain* of the integrals (6) are present, which necessitate a special investigation, and into which I will not enter at present. (It has to do with those matrix elements in the Kepler problem which, according to Heisenberg, correspond to the transition from one hyperbolic orbit to another.)

I have confined myself here to the case of classical mechanics, without magnetic fields, because the relativistic magnetic generalisation does not seem to me to be sufficiently clear yet. But we can scarcely doubt that the complete parallel between the two new quantum theories will still stand when this generalisation is obtained.

We conclude with a general observation on the whole formal apparatus of §§ 2, 3, and 4. The basic orthogonal system was regarded as an absolutely *discrete* system of functions. Now, in the most important applications this is *not* the case. Not only in the hydrogen atom but also in heavier atoms the wave equation (31) must possess a continuous proper value spectrum as well as a line spectrum. The former manifests itself, for example, in the continuous *optical* spectra which adjoin the limit of the series. It appeared better, provisionally, not to burden the formulae and the line of thought with this generalisation, though it is indeed indispensable. The chief aim of this paper is to work out, in the clearest manner possible, the formal connection between the two theories, and this is certainly not changed, in any essential point, by the appearance of a continuous spectrum. An important precaution that we have always observed is not to postulate, without further investigation, the convergence of the development in a series of proper functions. This precaution is especially demanded by the *accumulation of the proper values at a finite point* (viz. the limit of the series). This accumulation is most intimately connected with the appearance of the continuous spectrum.

§ 5. Comparison of the Two Theories. Prospect of a Classical Understanding of the Intensity and Polarisation of the Emitted Radiation

If the two theories—I might reasonably have used the singular—should [1] be tenable in the form just given, *i.e.* for more complicated systems as well, then every discussion of the superiority of the one over the other has only an illusory object, in a certain sense. For they are completely equivalent from the mathematical point of view, and it can only be a question of the subordinate point of convenience of calculation.

[1] There is a special reason for leaving this question open. The two theories initially take the energy function over from ordinary mechanics. Now in the cases treated the *potential* energy arises from the interaction of particles, of which perhaps *one*, at least, may be regarded in wave mechanics also as forming a point, on account of its great mass (*cf.* A. Einstein, *Berl. Ber.*, 1925, p. 10). We must take into account the possibility that it is no longer permissible to take over from ordinary mechanics the statement for the potential energy, if *both* " point charges " are really extended states of vibration, which penetrate each other.

To-day there are not a few physicists who, like Kirchhoff and Mach, regard the task of physical theory as being merely a mathematical description (*as economical as possible*) of the empirical connections between observable quantities, *i.e.* a description which reproduces the connection, as far as possible, without the intervention of unobservable elements. On this view, mathematical equivalence has almost the same meaning as physical equivalence. In the present case there might perhaps appear to be a certain superiority in the matrix representation because, through its stifling of intuition, it does not tempt us to form space-time pictures of atomic processes, which must perhaps remain uncontrollable. In this connection, however, the following *supplement* to the proof of equivalence given above is interesting. The equivalence *actually* exists, and it also exists *conversely*. Not only can the matrices be constructed from the proper functions as shown above, but also, conversely, the functions can be constructed from the numerically given matrices. Thus the functions do not form, as it were, an *arbitrary* and *special* "fleshly clothing" for the bare matrix skeleton, provided to pander to the need for intuitiveness. This really would establish the superiority of the matrices, from the epistemological point of view. We suppose that in the equations

$$q_i^{ik} = \int u_i(x)u_k(x)dx \tag{33}$$

the *left-hand* sides are given numerically and the functions $u_i(x)$ are to be found. (*N.B.*—The "density function" is omitted for simplicity; the $u_i(x)$'s *themselves* are to be orthogonal functions for the present.) We may then calculate by matrix multiplication (without, by the way, any "revolving", *i.e.* integration by parts) the following integrals,

$$\int P(x)u_i(x)u_k(x)dx, \tag{34}$$

where $P(x)$ signifies *any* power product of the q_l's. The totality of these integrals, when i and k are fixed, forms what is called the totality of the "*moments*" of the function $u_i(x)u_k(x)$. And it is known that, under very general assumptions, a function is determined uniquely by the totality of its moments. So all the products $u_i(x)u_k(x)$ are uniquely fixed, and thus also the squares $u_i(x)^2$, and therefore also $u_i(x)$ itself. The only arbitrariness lies in the supplementary detachment of the density function $\rho(x)$, *e.g.* $r^2 \sin\theta$ in polar co-ordinates. No false step is to be feared there, certainly not so far as *epistemology* is concerned.

Moreover, the validity of the thesis that mathematical and physical equivalence mean the same thing, must itself be qualified. Let us think, for example, of the two expressions for the electrostatic energy of a system of charged conductors, the space integral $\frac{1}{2}\int E^2 d\tau$ and the sum $\frac{1}{2}\Sigma e_i V_i$ taken over the conductors. The two expressions

are completely equivalent in electrostatics; the one may be derived from the other by integration by parts. Nevertheless we intentionally prefer the first and say that *it* correctly localises the energy in space. In the domain of electrostatics this preference has admittedly no justification. On the contrary, it is due simply to the fact that the first expression remains useful in electrodynamics also, while the second does not.

We cannot yet say with certainty to which of the two new quantum theories preference should be given, from *this* point of view. As the natural advocate of one of them, I will not be blamed if I frankly—and perhaps not wholly impartially—bring forward the arguments in its favour.

Leaving aside the special optical questions, the problems which the course of development of atomic dynamics brings up for consideration are presented to us by experimental physics in an eminently intuitive form; as, for example, how two colliding atoms or molecules rebound from one another, or how an electron or α-particle is diverted, when it is shot through an atom with a given velocity and with the initial path at a given perpendicular distance from the nucleus. In order to treat such problems more particularly, it is necessary to survey clearly the transition between macroscopic, perceptual mechanics and the micro-mechanics of the atom. I have lately [1] explained how I picture this transition. Micro-mechanics appears as a refinement of macro-mechanics, which is necessitated by the geometrical and mechanical smallness of the objects, and the transition is of the same nature as that from geometrical to physical optics. The latter is demanded as soon as the wave length is no longer very great compared with the dimensions of the objects investigated or with the dimensions of the space inside which we wish to obtain more accurate information about the light distribution. To me it seems extraordinarily difficult to tackle problems of the above kind, as long as we feel obliged on epistemological grounds to repress intuition in atomic dynamics, and to operate only with such abstract ideas as transition probabilities, energy levels, etc.

An especially important question—perhaps the cardinal question of all atomic dynamics—is, as we know, that of the *coupling* between the dynamic process in the atom and the electromagnetic field, or whatever has to appear in the place of the latter. Not only is there connected with this the whole complex of questions of dispersion, of resonance- and secondary-radiation, and of the natural breadth of lines, but, in addition, the specification of certain quantities in atomic dynamics, such as emission frequencies, line intensities, etc., has only a mere dogmatic meaning until this coupling is described mathematically in some form or other. Here, now, the matrix representation of atomic dynamics has led to the conjecture that in fact the electromagnetic field also *must* be represented otherwise, namely, by matrices, so that the coupling may be mathematically formulated. Wave mechanics

[1] Part II.

shows we are not compelled to do this in any case, for the mechanical field scalar (which I denote by ψ) is perfectly capable of entering into the unchanged Maxwell-Lorentz equations between the electromagnetic field vectors, as the "source" of the latter; just as, conversely, the electrodynamic potentials enter into the coefficients of the wave equation, which defines the field scalar.[1] In any case, it is worth while *attempting* the representation of the coupling in such a way that we bring into the unchanged Maxwell-Lorentz equations as *four-current* a four-dimensional vector, which has been suitably derived from the mechanical field scalar of the electronic motion (perhaps through the medium of the field vectors themselves, or the potentials). There even exists a hope that we can represent the wave equation for ψ equally well as a consequence of the Maxwell-Lorentz equations, namely, as an equation of continuity for electricity. The difficulty in regard to the problem of *several* electrons, which mainly lies in the fact that ψ is a function in *configuration* space, not in real space, must be mentioned. Nevertheless I would like to discuss the one-electron problem a little further, showing that it may be possible to give an extraordinarily clear interpretation of intensity and polarisation of radiation in this manner.

Let us consider the picture, on the wave theory, of the hydrogen atom, when it is in such a state that the field scalar ψ is given by a series of discrete proper functions, thus:

$$\psi = \sum_k c_k u_k(x) e^{\frac{2\pi\sqrt{-1}}{h} E_k t} \tag{35}$$

(x stands here for *three* variables, *e.g.* r, θ, ϕ; the c_k's are taken as real and it is correct to take the real part). We now make the *assumption* that the space density of electricity is given by the real part of

$$\psi \frac{\partial \bar{\psi}}{\partial t}. \tag{36}$$

The bar is to denote the conjugate complex function. We then calculate for the space density,

$$\text{space density} = 2\pi \sum_{(k,m)} c_k c_m \frac{E_k - E_m}{h} u_k(x) u_m(x) \sin \frac{2\pi t}{h}(E_m - E_k), \tag{37}$$

where the sum is to be taken once only over every combination (k, m). Only term *differences* enter (37) as frequencies. The former are so low that the length of the corresponding ether wave is large compared

[1] Similar ideas are expressed by K. Lanczos in an interesting note that has just appeared (*Ztschr. f. Phys.* 35, p. 812, 1926). This note is also valuable as showing that Heisenberg's atomic dynamics is capable of a continuous interpretation as well. However, Lanczos' work has fewer points of contact with the present work than at first it was thought to have. The determination of his formal system, which was provisionally left quite indefinite, is *not* to be sought by following the idea that in some way the symmetrical nucleus $K(s, \sigma)$ of Lanczos can be identified with the *Green's function* of our wave equation (21) or (31). For this Green's function, if it exists, has the quantum levels themselves as proper values. On the other hand, it is required that Lanczos' function should have the *reciprocals* of the quantum levels as proper values.

with atomic dimensions, that is, compared with the region within which (37) is markedly different from zero.[1] The radiation can therefore be estimated simply by the *dipole moment* which according to (37) the whole atom possesses. We multiply (37) by a Cartesian co-ordinate q_l, and by the "density function" $\rho(x)$, ($r^2 \sin\theta$ in the present case) and integrate over the whole space. According to (13), we get for the component of the dipole moment in the direction q_l,

$$(38) \qquad Mq_l = 2\pi \sum_{(k,m)} c_k c_m q_l^{km} \frac{E_k - E_m}{h} \sin \frac{2\pi t}{h}(E_m - E_k).$$

Thus we really get a "Fourier development" of the atom's electric moment, in which only term *differences* appear as frequencies. The Heisenberg matrix elements q_l^{km} come into the coefficients in such a manner that their co-operating influence on the intensity and polarisation of the part of the radiation concerned is completely intelligible on the grounds of classical electrodynamics.

The present sketch of the mechanism of radiation is far from completely satisfactory and is in no way final. Assumption (36) makes use, somewhat freely, of complex calculation, in order to put to one side undesired components of vibration whose radiation cannot be investigated at all in the simple way used for the dipole moment of the entire atom, because the corresponding ether wave lengths (about 0·01 Å) lie far below atomic dimensions. Moreover, if we integrate over all space, then by (5) the space density (37) gives zero and not, as is required, a finite value, independent of the time, which requires to be normalised to the electronic charge. In conclusion, for completeness, account should be taken of magnetic radiation, since if there is a spatial distribution of electric currents, radiation is possible without the appearance of an electric moment, *e.g.* with a frame aerial.

Nevertheless it appears to be a well-founded hope that a real understanding of the nature of emitted radiation will be obtained on the basis of one of the two very similar analytical mechanisms which have been sketched here.

(Received March 18, 1926).

[1] *Ann. d. Phys.* 79, p. 371, 1926, *i.e.* beginning of § 2, Part I. here.