"Zur Quantenmechanik," Zeit. Phys. **34** (1925), 858-888.

On Quantum Mechanics

By M. Born and P. Jordan in Göttingen

(Received on 27 September 1925)

Translated by D. H. Delphenich

The Ansätze that **Heisenberg** recently gave will be developed into a systematic theory of quantum mechanics (initially for systems of one degree of freedom). The mathematical tool is the matrix calculus. After it is briefly presented, the mechanical equations of motion will be derived from a variational principle, and it will be proved that the law of energy and **Bohr**'s frequency condition will follow from the mechanical equations on the grounds of **Heisenberg**'s quantum condition. The question of the uniqueness of the solution and the meaning of the phases in the partial oscillations will be discussed in the example of the anharmonic oscillator. The conclusion includes an attempt to introduce the laws of electromagnetic fields into the new theory.

Introduction. – The Ansätze that **Heisenberg** ([1]) recently published in this Zeitschrift for a new kinematics and mechanics that would correspond to the basic demands of quantum theory seem, to us, to have a broader scope. They signify an attempt to justify the new facts by the creation of a new, actually-measurable system of concepts, rather than a more-or-less heuristic and forced adaptation of the old known concepts. **Heisenberg** has expressed the physical concepts that guided him so clearly that any extended remark would seem superfluous. However, in a formal, mathematical context, his considerations are still in an early stage of development, as he himself emphasized. He explained his hypotheses only in simple examples and did not advance to a general theory. Encouraged by the fact that we can already understand his argument *in statu nascendi*, after he had completed his investigations, we endeavored to clarify the formal mathematical content of his Ansätzen, and some of our results will be presented here. They show that it is actually possible to erect the structure of a closed mathematical theory of quantum mechanics on the foundations that **Heisenberg** gave that has a remarkably close analogy with classical mechanics, but which still respects the features that characterize quantum phenomena.

With **Heisenberg**, we initially restrict ourselves to systems with *one degree of freedom*, which we assume to be *periodic* (classical speaking). We will address the generalization of the mathematical theory to systems of arbitrarily many degrees of freedom, as well as to aperiodic motions, in a continuation of this treatise. As an essential generalization of the **Heisenberg** Ansätze, we will not restrict our treatment to either non-relativistic mechanics or to calculations with Cartesian coordinates. The

---

([1]) **W. Heisenberg**, Zeit. Phys. **33** (1925), 879.

single restriction that we will impose in regard to coordinates lies in the fact our considerations will refer to *libration coordinates*, which are *periodic* functions of time in the classical theory. In general, in many cases, it seems natural to employ other coordinates – for example, the rotation angle φ for the rotator, which is a linear function of time. **Heisenberg** also proceeded in that way in his treatment of the rotator. However, it must remain undecided whether the processes that he applied there can be justified from the standpoint of a consistent theory of quantum mechanics.

The mathematical basis for **Heisenberg**'s consideration is the *multiplication law* for quantum-theoretic quantities, which he deduced by a clever analogy. The depiction of his formalism that we shall give here rests upon the remark that this rule is nothing but the law of *matrix multiplication* that is well-known to mathematicians. A square array (with discrete or continuous indices) that is infinite in both directions – viz., a so-called *matrix* – is the representative of a physical quantity that is given as a function of time in the classical theory. The mathematical model for the new quantum mechanics is then characterized by the use of *matrix analysis*, in place of the usual numerical analysis.

We have attempted to touch upon the simplest questions of mechanics and electrodynamics with those methods here. Reasoning by analogy, a natural *variational principle* will yield *equations of motion* for the most general **Hamiltonian** function in a close analogy with the classical canonical equations. The quantum condition, combined with a relation that flows out of the equations of motion, admits a simple matrix notation. With its help, one will succeed in proving the general validity of the *law of energy* and **Bohr**'s *frequency condition* in the sense that **Heisenberg** assumed, which is a proof that he also could not complete for the simple examples that be treated. We shall then return to one of those examples in more detail in order to get some indication of the role that the phases of the partial oscillations play in the new theory. In conclusion, we will show that the basic laws of the electromagnetic field in vacuum can also be easily addressed by the new method, and that will give a basis for the assumption that **Heisenberg** made that the squares of the contributions of the elements of the matrix that represents the electric moment of an atom gives a measure of the transition probabilities.

Chapter I – Matrix analysis

§ 1. Elementary operations. Functions. – We shall calculate with infinite square *matrices* ([1]), which we would like to denote by bold-faced symbols here, while the conventional symbols shall always mean ordinary numbers:

$$\mathbf{a} = (a\,(n\,m)) = \begin{bmatrix} a(00) & a(01) & a(02) & \cdots \\ a(10) & a(11) & a(12) & \cdots \\ a(20) & a(21) & a(22) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

([1]) One can find more details on matrix analysis in, say, **M. Bôcher**, *Einführung in die höhere Algebra*; in English, the translation by **Hans Beck**, Leipzig, Teubner, 1910, §§ 22 to 23. Furthermore, one can confer **R. Courant** and **D. Hilbert**, *Methoden der mathematischen Physik I*, Berlin, Springer, 1924; Chap. 1.

The equality of two matrices means the equality of the corresponding components:

$$\mathbf{a} = \mathbf{b} \qquad \text{means} \qquad a\,(nm) = b\,(nm). \tag{1}$$

Addition will be defined by the addition of corresponding components:

$$\mathbf{a} = \mathbf{b} + \mathbf{c} \qquad \text{means} \qquad a\,(nm) = b\,(nm) + c\,(nm). \tag{2}$$

Multiplication is defined by the "rows times columns" rule that is known from the theory of determinants:

$$\mathbf{a} = \mathbf{bc} \qquad \text{means} \qquad a\,(nm) = \sum_{k=0}^{\infty} b(nk)\,c(km). \tag{3}$$

Powers are defined by repeated multiplication. The associative law for multiplication and the distributive one for the combination of addition and multiplication are both true:

$$(\mathbf{ab})\,\mathbf{c} = \mathbf{a}\,(\mathbf{bc}), \tag{4}$$

$$\mathbf{a}\,(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac}. \tag{5}$$

By contrast, the commutative law of multiplication is *not* true: The equation $\mathbf{ab} = \mathbf{ba}$ is not generally correct. If it were true then **a** and **b** would be said to *commute*. The *identity matrix* that is defined by:

$$\mathbf{1} = (\delta_{mn}) \qquad \begin{cases} \delta_{nm} = 0 & \text{for} \quad n \neq m \\ \delta_{nn} = 1 \end{cases} \tag{6}$$

has the property that:

$$\mathbf{a1} = \mathbf{1a} = \mathbf{a}. \tag{6a}$$

The matrix $\mathbf{a}^{-1}$ that is *reciprocal* to **a** is defined by ([1]):

$$\mathbf{a}^{-1}\,\mathbf{a} = \mathbf{a}\,\mathbf{a}^{-1} = \mathbf{1}. \tag{7}$$

We say the *mean value* of a matrix **a** to mean the matrix whose diagonal elements coincide with those of **a**, while all of the remaining elements are zero:

$$\overline{\mathbf{a}} = (\delta_{nm}\,a\,(nm)). \tag{8}$$

The *diagonal sum* of the matrix **a** shall mean the sum of these diagonal elements, and $D\,(\mathbf{a})$ will denote:

$$D\,(\mathbf{a}) = \sum_{n} a\,(nn)\,. \tag{9}$$

([1]) It is known for *finite* square matrices that $\mathbf{a}^{-1}$ is always established uniquely by this definition when the *determinant* A of **a** is non-zero. If $A = 0$ then there is no reciprocal matrix to **a**.

One easily proves from (3) that: If the diagonal sum of a product $\mathbf{y} = \mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_m$ is finite then it will remain unchanged under cyclic permutation of the factors:

$$D\ (\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_m) = D\ (\mathbf{x}_r\ \mathbf{x}_{r+1}\ \ldots\ \mathbf{x}_m\ \mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_{r-1}). \tag{10}$$

It is obviously sufficient to convince oneself of the validity of theorem for *two* factors.

If the components of the matrices **a**, **b** are functions of one parameter *t* then:

$$\frac{d}{dt}\sum_k a(nk)\,b(km) = \sum_k \{\dot{a}(nk)\,b(km) + a(nk)\,\dot{b}(km)\},$$

or, from the definition (3):

$$\frac{d}{dt}(\mathbf{ab}) = \dot{\mathbf{a}}\mathbf{b} + \mathbf{a}\dot{\mathbf{b}}\,. \tag{11}$$

A repeated application of (11) gives:

$$\frac{d}{dt}(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_n) = \dot{\mathbf{x}}_1\mathbf{x}_2\cdots\mathbf{x}_n + \mathbf{x}_1\dot{\mathbf{x}}_2\cdots\mathbf{x}_n + \cdots + \mathbf{x}_1\mathbf{x}_2\cdots\dot{\mathbf{x}}_n\,. \tag{11′}$$

Functions of matrices will be defined by the calculation processes (2), (3). Initially, the most general function $\mathbf{f}\ (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ to come under consideration here shall be one that can be formally represented by a sum of finitely or infinitely-many products of powers in the arguments $\mathbf{x}_k$ with *numbers* as coefficients. Functions $\mathbf{y}_l\ (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ can also be defined then by equations:

$$\left.\begin{aligned} &\mathbf{f}_1(\mathbf{y}_1,\ldots,\mathbf{y}_n;\mathbf{x}_1,\ldots,\mathbf{x}_n) = 0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &\mathbf{f}_n(\mathbf{y}_1,\ldots,\mathbf{y}_n;\mathbf{x}_1,\ldots,\mathbf{x}_n) = 0. \end{aligned}\right\} \tag{12}$$

Namely, in order to obtain functions $\mathbf{y}_l$ of the aforementioned form that satisfy equations (12), one needs only to develop the $\mathbf{y}_l$ as series in increasing powers of the $\mathbf{x}_k$ and determine the coefficients of the series by substituting the former series into (12). One sees that this will always imply just as many equations as unknowns. The number of equations and unknowns is admittedly larger than it is in the application of the method of undetermined coefficients in the usual analysis that calculates with *commutative* multiplication. Upon substituting the series for the $\mathbf{y}_l$ in each of the equations (12) and combining the associated terms one will obtain summands of the form $C'\mathbf{x}_1\ \mathbf{x}_2$, as well as ones of the form $C''\mathbf{x}_2\ \mathbf{x}_1$, and C', as well as C'' (and not just $C' + C''$), must vanish separately. However, two terms $\mathbf{x}_1\ \mathbf{x}_2$ and $\mathbf{x}_2\ \mathbf{x}_1$ with two available coefficients will also appear in the development of each $\mathbf{y}_l$ then.

§ 2. Symbolic differentiation. – A process of calculation that will be much more useful later on that we would like to consider in detail here shall be referred to as the *differentiation* of a matrix function. However, one should note that this process possesses properties that are similar to those of the differentiation in ordinary analysis only in some respects. For example, the product rule of differentiation and the rule for the differentiation of a function of a function will no longer be valid, in general. Only when all of the matrices that appear *commute* with each other will all of the rules of ordinary analysis be valid for this differentiation.

Let:

$$\mathbf{y} = \prod_{m=1}^{s} \mathbf{x}_{l_m} = \mathbf{x}_{l_1} \mathbf{x}_{l_2} \cdots \mathbf{x}_{l_s} . \tag{13}$$

We define:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}_k} = \sum_{r=1}^{s} \delta_{l_r k} \prod_{m=r+1}^{s} \mathbf{x}_{l_m} \prod_{m=1}^{r-1} \mathbf{x}_{l_m} \qquad \begin{cases} \delta_{jk} = 0 & \text{for} \quad j \neq k, \\ \delta_{kk} = 1. \end{cases} \tag{14}$$

In words, that rule reads: Think of all factors in the given product as being written out *individually* (so, e.g., $\mathbf{x}_1\ \mathbf{x}_1\ \mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_2$, instead of $\mathbf{x}_1^3\,\mathbf{x}_2^2$). Remove any factor $\mathbf{x}_k$ and define the product of all factors that follow it and all factors that precede it (in that sequence). The sum of all terms that are constructed in that way will be the differential quotient of the product with respect to that $\mathbf{x}_k$.

Some examples might clarify this process:

$$\mathbf{y} = \mathbf{x}^n, \qquad \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = n\,\mathbf{x}^{n-1},$$

$$\mathbf{y} = \mathbf{x}_1^n\,\mathbf{x}_2^m, \qquad \frac{\partial \mathbf{y}}{\partial \mathbf{x}_1} = \mathbf{x}_1^{n-1}\,\mathbf{x}_2^m + \mathbf{x}_1^{n-2}\,\mathbf{x}_2^m\,\mathbf{x}_1 + \cdots + \mathbf{x}_2^m\,\mathbf{x}_1^{n-1},$$

$$\mathbf{y} = \mathbf{x}_1^2\,\mathbf{x}_2\,\mathbf{x}_1\,\mathbf{x}_3, \qquad \frac{\partial \mathbf{y}}{\partial \mathbf{x}_1} = \mathbf{x}_1\,\mathbf{x}_2\,\mathbf{x}_1\,\mathbf{x}_3 + \mathbf{x}_2\,\mathbf{x}_1\,\mathbf{x}_3\,\mathbf{x}_1 + \mathbf{x}_3\,\mathbf{x}_1^2\,\mathbf{x}_2 .$$

Furthermore, if we demand that:

$$\frac{\partial(\mathbf{y}_1 + \mathbf{y}_2)}{\partial \mathbf{x}_k} = \frac{\partial \mathbf{y}_1}{\partial \mathbf{x}_k} + \frac{\partial \mathbf{y}_2}{\partial \mathbf{x}_k} \tag{15}$$

then the derivative $\partial\mathbf{y} / \partial\mathbf{x}$ will be defined for the most general function $\mathbf{y}$.

With those definitions and that of the diagonal sum (9), one has the relation:

$$\frac{\partial D(\mathbf{y})}{\partial x_k(nm)} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}_k}(mn), \tag{16}$$

in which the *mn*-component of the matrix $\partial\mathbf{y} / \partial\mathbf{x}_k$ is on the right-hand side. In order to prove (16), it is obviously sufficient to consider a function **y** of the form (13). From (14) and (3), one will have:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}_k}(mn) = \sum_{r=1}^{s} \delta_{l_r k} \sum_{\tau} \prod_{p=r+1}^{s} x_{l_p}(\tau_p \tau_{p+1}) \prod_{p=1}^{r-1} x_{l_p}(\tau_p \tau_{p+1}); \tag{17}$$

$$\tau_{r+1} = m, \qquad \tau_{r+1} = \tau_1, \qquad \tau_r = n.$$

On the other hand, one can infer from (3) and (9) that:

$$\frac{\partial D(\mathbf{y})}{\partial x_k(nm)} = \sum_{r=1}^{s} \delta_{l_r k} \sum_{\tau} \prod_{p=1}^{r-1} x_{l_p}(\tau_p \tau_{p+1}) \prod_{p=r+1}^{s} x_{l_p}(\tau_p \tau_{p+1}), \tag{17′}$$

$$\tau_1 = \tau_{s+1}, \qquad \tau_r = n, \qquad \tau_{r+1} = m.$$

A comparison of (17) and (17′) will give (16).

Let us emphasize a fact right now that will be important later and which can be read off from the definition (14):

The partial derivatives of a product are invariant under cyclic permutations of the factors.

This theorem will also follow from (10), due to (16).

To conclude these preliminary remarks, a few words shall be devoted to the functions **g** (**p**, **q**) of *two* variables. For:

$$\mathbf{y} = \mathbf{p}^s \mathbf{q}^r, \tag{18}$$

one will have, from (14):

$$\frac{\partial \mathbf{y}}{\partial \mathbf{p}} = \sum_{l=0}^{s-1} \mathbf{p}^{s-1-l} \mathbf{q}^r \mathbf{p}^l, \qquad \frac{\partial \mathbf{y}}{\partial \mathbf{q}} = \sum_{j=0}^{s-1} \mathbf{q}^{r-1-j} \mathbf{p}^s \mathbf{q}^j. \tag{18′}$$

From § **1**, the most general function **g** (**p**, **q**) to be considered can be represented by a linear combination of terms of the form:

$$\mathbf{z} = \prod_{j=1}^{k} (\mathbf{p}^{s_j} \mathbf{q}^{r_j}). \tag{19}$$

With the abbreviation:

$$\mathbf{P}_l = \prod_{j=l+1}^{k} (\mathbf{p}^{s_j} \mathbf{q}^{r_j}) \prod_{j=1}^{l-1} (\mathbf{p}^{s_j} \mathbf{q}^{r_j}), \tag{20}$$

the derivatives can be written:

$$\left.\begin{aligned}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} &= \sum_{l=1}^{k}\sum_{m=0}^{s_l-1}\mathbf{p}^{s_l-1-m}\mathbf{q}^{r_l}\,\mathbf{P}_l\,\mathbf{p}^m,\\ \frac{\partial \mathbf{z}}{\partial \mathbf{q}} &= \sum_{l=1}^{k}\sum_{m=0}^{s_l-1}\mathbf{q}^{r_l-1-m}\,\mathbf{P}_l\,\mathbf{p}^{s_l}\mathbf{q}^m.\end{aligned}\right\} \tag{21}$$

We can infer an important consequence of these equations. We consider the matrices:

$$\mathbf{d}_1 = \mathbf{q}\frac{\partial \mathbf{z}}{\partial \mathbf{q}} - \frac{\partial \mathbf{z}}{\partial \mathbf{q}}\mathbf{q}, \qquad \mathbf{d}_2 = \mathbf{p}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} - \frac{\partial \mathbf{z}}{\partial \mathbf{p}}\mathbf{p}. \tag{22}$$

From (21), one will have:

$$\mathbf{d}_1 = \sum_{l=0}^{k}(\mathbf{q}^{r_l}\,\mathbf{P}_l\,\mathbf{p}^{s_l} - \mathbf{P}_l\,\mathbf{p}^{s_l}\mathbf{q}^{r_l}),$$

$$\mathbf{d}_2 = \sum_{l=0}^{k}(\mathbf{p}^{s_l}\,\mathbf{q}^{r_l}\,\mathbf{P}_l - \mathbf{q}^{r_l}\,\mathbf{P}_l\,\mathbf{p}^{s_l}),$$

and it will follow from this that:

$$\mathbf{d}_1 + \mathbf{d}_2 = \sum_{l=0}^{k}(\mathbf{p}^{s_l}\,\mathbf{q}^{r_l}\,\mathbf{P}_l - \mathbf{P}_l\,\mathbf{p}^{s_l}\mathbf{q}^{r_l}).$$

The second term of one summand in this will always cancel the first term of the following one, and the first and last terms will cancel the entire sum. Hence:

$$\mathbf{d}_1 + \mathbf{d}_2 = 0. \tag{23}$$

Due to its linear character in **z**, this relation is true for not only the expressions **z** of the form (19), but, at the same time, also for arbitrary analytic functions **g** (**p**, **q**) ([1]).

To conclude this brief presentation of matrix analysis, we would like to prove the theorem:

Any matrix equation:

$$\mathbf{F}\,(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r) = 0$$

will remain true when one subjects all argument matrices $\mathbf{x}_j$ *to one and the same permutation of all rows and columns.*

It is obviously enough to show that for the two matrices **a**, **b** that go to **a′**, **b′** under that operation, the invariances:

([1]) More generally, for functions of *r* variables, one will have:

$$\sum_r\left(\mathbf{x}_r\frac{\partial \mathbf{g}}{\partial \mathbf{x}_r} - \frac{\partial \mathbf{g}}{\partial \mathbf{x}_r}\mathbf{x}_r\right) = 0.$$

$$\mathbf{a}' + \mathbf{b}' = (\mathbf{a} + \mathbf{b})', \qquad \mathbf{a}'\,\mathbf{b}' = (\mathbf{a}\,\mathbf{b})'$$

will be true, in which the right-hand side means those matrices that arise from **a** + **b** and **ab**, resp., by those permutations.

We shall prove this when we replace the operation of permutation with multiplication by a suitable matrix ([1]).

We write a permutation as:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & \cdots \\ k_0 & k_1 & k_2 & k_3 & \cdots \end{pmatrix} = \begin{pmatrix} n \\ k_n \end{pmatrix}.$$

We associate this with the *permutation matrix:*

$$\mathbf{p} = (p\,(nm)), \qquad p(nm) = \begin{cases} 1 & \text{for } m = k_n, \\ 0 & \text{otherwise.} \end{cases}$$

Let the transpose of the matrix **p** be:

$$\tilde{\mathbf{p}} = (\tilde{p}\,(nm)), \qquad \tilde{p}\,(nm) = \begin{cases} 1 & \text{for } n = k_n, \\ 0 & \text{otherwise.} \end{cases}$$

Upon multiplying both of these, it will follow that:

$$\mathbf{p}\,\tilde{\mathbf{p}} = \left(\sum_k p\,(nk)\,\tilde{p}\,(km) \right) = (\delta_{nm}) = \mathbf{1},$$

since both factors $p\,(nk)$ and $\tilde{p}\,(km)$ are simultaneously non-zero only when $k = k_n = k_m$, so $n = m$. Therefore, $\tilde{\mathbf{p}}$ is the reciprocal of **p**:

$$\tilde{\mathbf{p}} = \mathbf{p}^{-1}.$$

Let **a** be an arbitrary matrix, so:

$$\mathbf{p}\,\mathbf{a} = \left(\sum_k p\,(nk)\,a\,(km) \right) = (a\,(k_n, m))$$

([1]) The process of proof that is chosen here has the advantage that it can make the close connection between permutations and an important class of general matrix transformations more clear. However, the validity of the theorem in question can also be inferred directly from the remark that in the definitions of *equality*, as well as *addition* and *multiplication*, of matrices, no use was made of the ordering of the rows (columns, resp.).

will be a matrix that arises from **a** by a permutation $\begin{pmatrix} n \\ k_n \end{pmatrix}$ of the rows. If one and the same permutation is applied to rows and columns then that will yield the matrix:

$$\mathbf{a}' = \mathbf{p}\,\mathbf{a}\,\mathbf{p}^{-1}.$$

It follows from this with nothing further that:

$$\begin{aligned} \mathbf{a}' + \mathbf{b}' &= \mathbf{p}\,(\mathbf{a}+\mathbf{b})\,\mathbf{p}^{-1} = (\mathbf{a}+\mathbf{b})', \\ \mathbf{a}'\,\mathbf{b}' &= \mathbf{p}\,\mathbf{a}\,\mathbf{b}\,\mathbf{p}^{-1} = (\mathbf{a}\,\mathbf{b})', \end{aligned}$$

with which our assertion is proved.

One then sees that no sequence or ordering of the elements can be determined by matrix equations.

Moreover, the much more general theorem is obviously true that any matrix equation is invariant under transformations of the form:

$$\mathbf{a}' = \mathbf{b}\,\mathbf{a}\,\mathbf{b}^{-1},$$

in which **b** means an *arbitrary* matrix. Of course, later on, we will see that this is no longer true for matrix differential equations with no further assumptions.

Chapter II – Dynamics

§ 3. The basic laws. – The dynamical system is described by *coordinates* **q** and *impulses* **p**. They shall be represented by matrices:

$$\mathbf{q} = (q\,(nm)\,e^{2\pi i r\,\nu(nm)\,t}), \qquad \mathbf{p} = (p\,(nm)\,e^{2\pi i r\,\nu(nm)\,t})\,. \tag{24}$$

The $\nu\,(nm)$ in these expressions mean the quantum-theoretical frequencies that belong to the transitions between the states with *quantum numbers n* and *m*. The matrices (24) shall be *Hermitian;* i.e., under matrix transposition, each component shall go to its conjugate value, and indeed that must be true for all real *t*. We will then have:

$$q(nm)\,q\,(mn) = |\,q\,(nm)\,|^2 \tag{25}$$

and

$$\nu\,(nm) = -\,\nu\,(mn). \tag{26}$$

If **q** is a *Cartesian* coordinate then the quantity (25) will be definitive for the *probabilities* ([1]) of the transitions $n \rightleftarrows m$.

We would like to further demand that:

([1]) On this subject, see § **8**.

$$\nu(nm) + \nu(kl) + \nu(lj) = 0. \tag{27}$$

With (26), that be expressed as: There are quantities W_n such that:

$$h\nu(nm) = W_n - W_m. \tag{28}$$

It follows from this, with (2), (3), that the function $\mathbf{g}(\mathbf{p}, \mathbf{q})$ will always take the form:

$$\mathbf{g} = (g(nm)\, e^{2\pi i\, \nu(nm)\, t})$$

again, and indeed, the matrix $(g(mn))$ will emerge in that way from the matrices $(q(nm))$, $(p(nm))$ by just the same process by which $\mathbf{g}$ was obtained from $\mathbf{q}$, $\mathbf{p}$. For that reason, from now on, we can choose the briefer notation:

$$\mathbf{q} = (q(nm)), \qquad \mathbf{p} = (p(nm)), \tag{30}$$

in place of the representation (24), which we shall abandon.

When we once more recall (24), [(29), resp.], we will get the matrix:

$$\dot{\mathbf{g}} = 2\pi i\, (\nu(nm)\, g(nm)) \tag{31}$$

for the *temporal derivative* of the matrix $\mathbf{g} = (g(nm))$.

If $\nu(nm) \neq 0$ for $n \neq m$ (as we would like to assume) then $\dot{\mathbf{g}} = 0$ would mean that $\mathbf{g}$ is a diagonal matrix with $g(nm) = \delta_{nm}\, g(nm)$.

A differential equation $\dot{\mathbf{g}} = \mathbf{a}$ is invariant under the process that subjects the rows and columns of all matrices, as well as the numbers W_n to the same permutation. In order to see that, we consider the diagonal matrix:

$$\mathbf{W} = (\delta_{nm}\, W_n);$$

we will then have:

$$\mathbf{Wg} = \left(\sum_k \delta_{nk} W_n\, g(km) \right) = (\delta_{nm}\, g(nm)),$$

$$\mathbf{gW} = \left(\sum_k g(nk)\, \delta_{km} W_k \right) = (W_m\, g(nm)),$$

so, from (31):

$$\dot{\mathbf{g}} = \frac{2\pi i}{h} ((W_n - W_m)\, g(nm)) = \frac{2\pi i}{h} (\mathbf{Wg} - \mathbf{gW}).$$

Now, if p is a permutation matrix then the transform:

$$\mathbf{W}' = \mathbf{p}\, \mathbf{W}\, \mathbf{p}^{-1} = (\delta_{n_k m}\, W_{n_k})$$

will be a diagonal matrix with the permuted W_n in the diagonal. One will then have:

$$\mathbf{p}\,\dot{\mathbf{g}}\,\mathbf{p}^{-1} = \frac{2\pi i}{h}(\mathbf{W}'\mathbf{g}' - \mathbf{g}'\mathbf{W}') = \dot{\mathbf{g}}',$$

in which $\mathbf{g}' = \mathbf{p}\,\mathbf{g}\,\mathbf{p}^{-1}$, and $\dot{\mathbf{g}}'$ means the temporal derivative of $\mathbf{g}'$ that is constructed using the rule (31) with permuted W_n .

The rows and columns of $\dot{\mathbf{g}}$ then suffer the same permutation as those of **g**, and our assertion follows from that.

It should be noted that a corresponding theorem for an arbitrary transformation of the form $\mathbf{a}' = \mathbf{b}\,\mathbf{a}\,\mathbf{b}^{-1}$ is not true, since $\mathbf{W}'$ would no longer be a diagonal matrix in that case. Despite that difficulty, a closer study of these general transformations seems imperative to us, because it promises to give a glimpse into the deeper connections of the new theory; we will come back to this later ([1]).

For the case of a **Hamiltonian** function of the form:

$$\mathbf{H} = \frac{1}{2m}\mathbf{p}^2 + \mathbf{U}(\mathbf{q}),$$

we will assume, with **Heisenberg**, that the *equations of motion* read just like the classical ones, such that we can write:

$$\left.\begin{aligned} \dot{\mathbf{q}} &= \frac{\partial \mathbf{H}}{\partial \mathbf{p}} = \frac{1}{m}\mathbf{p}, \\ \dot{\mathbf{p}} &= -\frac{\partial \mathbf{H}}{\partial \mathbf{q}} = -\frac{\partial \mathbf{U}}{\partial \mathbf{q}}, \end{aligned}\right\} \tag{32}$$

with the symbolism of § **2**.

We shall attempt to determine associated equations of motion for the general case of an arbitrary **Hamiltonian** function **H** (**p**, **q**), as well, by reasoning by analogy. That will be necessary when one goes on to relativistic mechanics, and especially to the treatment of the motion of electrons that interact with magnetic fields. In the latter case, the function **H** can no longer be represented in Cartesian coordinates as the sum of two functions, one of which depends upon only the impulses and the other of which depends upon only the coordinates.

Classically, the equations of motion are derived from the action principle:

$$\int_{t_0}^{t_1} L\,dt = \int_{t_0}^{t_1} \{p\dot{q} - H(p,q)\}dt \;= \text{extremum.} \tag{33}$$

If we think of inserting the Fourier development of L in this, and we take the time interval $t_1 - t_0$ to be sufficiently large then only the constant term of L will contribute to the integral. The form that the action principle will then take is then closely related to the following fact in quantum mechanics:

([1]) Cf., the continuation of this paper that is soon to appear.

The diagonal sum $D(\mathbf{L}) = \sum_k L(kk)$ *shall be extremized:*

$$D(\mathbf{L}) = D(\mathbf{p}\dot{\mathbf{q}} - \mathbf{H}(\mathbf{p},\mathbf{q})) = \text{extremum}, \tag{34}$$

and indeed for a suitable choice of **p** *and* **q** *and with given* $\nu(nm)$.

When one sets the derivatives of $D(\mathbf{L})$ with respect to the elements of **p** and **q** equal to zero, one will then have the equations of motion:

$$2\pi i\, \nu(nm)\, q(nm) = \frac{\partial D(\mathbf{L})}{\partial p(mn)},$$

$$2\pi i\, \nu(nm)\, p(nm) = \frac{\partial D(\mathbf{L})}{\partial q(mn)}.$$

One sees from (26), (31), and (16) that these equations of motion can generally be written in the *canonical* form:

$$\left.\begin{aligned} \dot{\mathbf{q}} &= \frac{\partial \mathbf{H}}{\partial \mathbf{p}}, \\ \dot{\mathbf{p}} &= -\frac{\partial \mathbf{H}}{\partial \mathbf{p}}. \end{aligned}\right\} \tag{35}$$

Heisenberg employed a relation that had been presented by **Thomas** ([1]) and **Kuhn** ([2]) as his quantum condition. The equation:

$$J = \oint p\, dq = \int_0^{1/\nu} p\, \dot{q}\, dt$$

of "classical" quantum theory can, when one appeals to the Fourier development of p and q:

$$p = \sum_{\tau=-\infty}^{\infty} p_\tau e^{2\pi i \nu \tau t}, \qquad q = \sum_{\tau=-\infty}^{\infty} q_\tau e^{2\pi i \nu \tau t},$$

be converted into:

$$1 = 2\pi i \sum_{\tau=-\infty}^{\infty} \tau \frac{\partial}{\partial J}(q_\tau, p_{-\tau}). \tag{36}$$

If $p = m\dot{q}$ in this then the p_τ can be expressed in terms of the q_τ, and one will get the classical equation whose analogous conversion into a differential equation will yield the

([1]) **W. Thomas**, Naturwiss. **13** (1925), 627.
([2]) **W. Kuhn**, Zeit. Phys. **33** (1925), 408.

relation of **Thomas** and **Kuhn**. Since the assumption $\mathbf{p} = m\dot{\mathbf{q}}$ should not be made here, we must immediately translate equation (36) into a differential equation.

One should have a correspondence between:

$$\sum_{\tau=-\infty}^{\infty} \tau \frac{\partial}{\partial J}(q_\tau, p_{-\tau}) \qquad \text{and} \qquad \frac{1}{h}\sum_{\tau=-\infty}^{\infty} (q(n+\tau,n)\,p(n,n+\tau) - q(n,n-\tau)\,p(n-\tau,n));$$

the q (nm), p (nm) on the right in this that include a negative index are set to zero. In that way, one will get the quantum condition:

$$\sum_k (p(nk)\,q(kn) - q(nk)\,p(kn)) = \frac{h}{2\pi i} \tag{37}$$

as the analogue of (36).

This represents an infinitude of equations, namely, one for each n.

In particular, for $\mathbf{p} = m\dot{\mathbf{q}}$, it will imply that:

$$\sum_k \nu(kn)\,|\,q(nk)\,|^2 = \frac{h}{8\pi^2 m},$$

which coincides with the **Heisenberg** form of the quantum condition – viz., the **Thomas-Kuhn** equation. (37) must be regarded as the proper generalization of that equation.

Moreover, one sees from (10) that the diagonal sum D (**p**, **q**) – D (**q**, **p**) = 0, while (37) leads to D (**p**, **q**) – D (**q**, **p**) = ∞. The matrices considered are therefore never finite ([1]).

§ 4. Consequences. Laws of energy and frequency. – With these preliminaries from the previous paragraphs, the basic laws of the new mechanics can be given completely. All further laws of quantum mechanics that should be endowed with general validity must be *provable* in terms of them. Some of the laws to be proved that mainly come under consideration are the *law of energy* and the ***Bohr*** *frequency condition.* The law of energy says that when **H** is the energy, one will have $\dot{\mathbf{H}} = 0$, or that **H** is a *diagonal matrix.* According to **Heisenberg**, the diagonal terms H (nn) of **H** will then be interpreted as the *energies of the various states of the system,* and the **Bohr** frequency condition will require that:

$$h\nu(nm) = H(nn) - H(mm),$$

or

$$W_n = H(nn) + \text{const.}$$

We consider the quantity:

$$\mathbf{d} = \mathbf{p}\,\mathbf{q} - \mathbf{q}\,\mathbf{p}.$$

From (11), (35) will become:

([1]) Nor do they belong to the class of "restricted" infinite matrices that have been considered almost exclusively by the mathematicians until now.

$$\dot{\mathbf{d}} = \dot{\mathbf{p}}\mathbf{q} + \mathbf{p}\dot{\mathbf{q}} - \dot{\mathbf{q}}\mathbf{p} - \mathbf{q}\dot{\mathbf{p}}$$

$$= \mathbf{q}\frac{\partial \mathbf{H}}{\partial \mathbf{q}} - \frac{\partial \mathbf{H}}{\partial \mathbf{q}}\mathbf{q} + \mathbf{p}\frac{\partial \mathbf{H}}{\partial \mathbf{p}} - \frac{\partial \mathbf{H}}{\partial \mathbf{p}}\mathbf{p}.$$

From (22), (23), one then has $\dot{\mathbf{d}} = 0$, and $\mathbf{d}$ is a diagonal matrix. However, the diagonal terms of $\mathbf{d}$ are established by the quantum condition (37) precisely. In summary, with the use of the identity matrix $\mathbf{1}$ that is defined by (6), we will get the equation:

$$\mathbf{p}\,\mathbf{q} - \mathbf{q}\,\mathbf{p} = \frac{h}{2\pi i}\mathbf{1}, \tag{38}$$

which we call the *sharpened quantum condition* and which we will base all further conclusions upon.

One can infer from the form of this equation that if an equation (*A*) were derived from (38) then (*A*) would remain true when one switched $\mathbf{p}$ with $\mathbf{q}$ and simultaneously replaced h with $-h$. For that reason, of the equations:

$$\mathbf{p}^n\,\mathbf{q} = \mathbf{q}\,\mathbf{p}^n + n\frac{h}{2\pi i}\mathbf{p}^{n-1}, \tag{39}$$

$$\mathbf{q}^n\,\mathbf{p} = \mathbf{p}\,\mathbf{q}^n - n\frac{h}{2\pi i}\mathbf{q}^{n-1}, \tag{39′}$$

only one them can be proved from (38), which can be done by induction quite easily.

We would now like to prove the laws of energy and frequency, as they were expressed above, and first for the case:

$$\mathbf{H} = \mathbf{H}_1(\mathbf{p}) + \mathbf{H}_2(\mathbf{q}).$$

From what was done in § **1**, $\mathbf{H}_1(\mathbf{p})$ and $\mathbf{H}_2(\mathbf{q})$ can be formally replaced by the power series:

$$\mathbf{H}_1(\mathbf{p}) = \sum_s a_s\,\mathbf{p}^n, \qquad \mathbf{H}_2(\mathbf{q}) = \sum_s b_s\,\mathbf{q}^n$$

in this. Formulas (39), (39′) then allow one to see that:

$$\left.\begin{aligned} \mathbf{H}\mathbf{q} - \mathbf{q}\mathbf{H} &= \frac{h}{2\pi i}\frac{\partial \mathbf{H}}{\partial \mathbf{p}}, \\ \mathbf{H}\mathbf{p} - \mathbf{p}\mathbf{H} &= -\frac{h}{2\pi i}\frac{\partial \mathbf{H}}{\partial \mathbf{q}}, \end{aligned}\right\} \tag{40}$$

and a comparison with the equations of motion (35) will yield:

$$\left.\begin{aligned} \dot{\mathbf{q}} &= \frac{2\pi i}{h}(\mathbf{Hq}-\mathbf{qH}), \\ \dot{\mathbf{p}} &= \frac{2\pi i}{h}(\mathbf{Hp}-\mathbf{pH}). \end{aligned}\right\} \tag{41}$$

Now, if the matrix **Hg** – **gH** is briefly denoted by $\left|\begin{matrix}\mathbf{H}\\ \mathbf{g}\end{matrix}\right|$ then one will have:

$$\left|\begin{matrix}\mathbf{H}\\ \mathbf{ab}\end{matrix}\right| = \left|\begin{matrix}\mathbf{H}\\ \mathbf{a}\end{matrix}\right|\mathbf{b}+\mathbf{a}\left|\begin{matrix}\mathbf{H}\\ \mathbf{b}\end{matrix}\right|;$$

however, for **g** = **g** (**p**, **q**), it will generally follow from this that:

$$\dot{\mathbf{g}} = \frac{2\pi i}{h}\left|\begin{matrix}\mathbf{H}\\ \mathbf{g}\end{matrix}\right| = \frac{2\pi i}{h}(\mathbf{Hg}-\mathbf{gH}). \tag{43}$$

For the proof, one then needs only to calculate $\dot{\mathbf{g}}$ as a function of **p**, **q** and $\dot{\mathbf{p}}$, $\dot{\mathbf{q}}$ by means of (11), (11), as well as $\left|\begin{matrix}\mathbf{H}\\ \mathbf{g}\end{matrix}\right|$ as a function of **p**, **q** and $\left|\begin{matrix}\mathbf{H}\\ \mathbf{p}\end{matrix}\right|$, $\left|\begin{matrix}\mathbf{H}\\ \mathbf{q}\end{matrix}\right|$, and to think of applying (41) then. If one sets **g** = **H** in (43), in particular, then one will get:

$$\dot{\mathbf{H}} = 0. \tag{44}$$

Once the law of energy has been proved and **H** is known as a diagonal matrix, (41) will take on the form:

$$h\,\nu(nm)\,q(nm) = (H(nn) - H(mm))\,q(nm),$$

$$h\,\nu(nm)\,p(nm) = (H(nn) - H(mm))\,p(nm),$$

from which, the frequency condition will follow.

If we now go on to the more general **Hamiltonian** functions $\mathbf{H}^* = \mathbf{H}^*(\mathbf{p}, \mathbf{q})$ then we will easily see from some examples, such as $\mathbf{H}^* = \mathbf{p}^2\,\mathbf{q}$, that we no longer have $\dot{\mathbf{H}}^* = 0$, in general. However, one sees that the **Hamiltonian** function $H = \frac{1}{2}(\mathbf{p}^2\,\mathbf{q} + \mathbf{q}\,\mathbf{p}^2)$ will produce the same equations of motion as $\mathbf{H}^*$, and that $\dot{\mathbf{H}}$ will once more be equal to zero. We then express the laws of energy and frequency as:

For every function $\mathbf{H}^* = \mathbf{H}^*(\mathbf{p}, \mathbf{q})$, *there is a function* $\mathbf{H} = \mathbf{H}(\mathbf{p}, \mathbf{q})$ *such that* $\mathbf{H}^*$ *and* **H** *will imply the same equations of motion when they are regarded as Hamiltonian functions, and that* **H** *will play the role of a temporally-constant energy that fulfills the frequency condition for these equations of motion.*

From the argument that was made above, it is sufficient that the given function **H** should satisfy equations (40), along with:

$$\frac{\partial \mathbf{H}}{\partial \mathbf{p}} = \frac{\partial \mathbf{H}^{\bullet}}{\partial \mathbf{p}}, \qquad \frac{\partial \mathbf{H}}{\partial \mathbf{q}} = \frac{\partial \mathbf{H}^{\bullet}}{\partial \mathbf{q}}. \tag{45}$$

From § **1**, $\mathbf{H}^*$ can be formally represented as a sum of products of powers of **p** and **q**, and due to the linearity of equations (40), (45) in **H**, $\mathbf{H}^*$, we will simply have to give the summand in **H** that corresponds to each summand in $\mathbf{H}^*$. We then need to consider only the case:

$$\mathbf{H}^* = \prod_{j=1}^{k} (\mathbf{p}^{s_j}\, \mathbf{q}^{r_j}). \tag{46}$$

From the remarks in § **2**, equations (45) must be fulfilled when **H** is set equal to a linear form in those products of powers of **p**, **q** that arise from $\mathbf{H}^*$ by cyclic permutation of the factors; in that way, only the sum of the coefficients will be taken to be equal to 1. It is not as easy to answer the question of how these coefficients must be chosen in order for equations (40) to be fulfilled. It might suffice here to treat the case of $k = 1$, so:

$$\mathbf{H}^* = \mathbf{p}^s\, \mathbf{q}^r. \tag{47}$$

Formula (39) can be generalized to ([1]):

$$\mathbf{p}^m\, \mathbf{q}^n - \mathbf{p}^n\, \mathbf{q}^m = m\frac{h}{2\pi i}\sum_{l=0}^{n-1} \mathbf{q}^{n-1-l}\, \mathbf{p}^{n-1}\, \mathbf{q}^l. \tag{48}$$

For $n = 1$, that is once more (39); in general, it will follow from (48) that due to (39), one will have:

$$\mathbf{p}^m\, \mathbf{q}^{n+1} - \mathbf{p}^{n+1}\, \mathbf{q}^m = (\mathbf{p}^m\, \mathbf{q}^n - \mathbf{p}^n\, \mathbf{q}^m)\, \mathbf{q} + m\frac{h}{2\pi i}\mathbf{q}^n\, \mathbf{p}^{m+1}.$$

Switching **p** and **q**, with a change of sign on *h*, will yield the new formula:

([1]) Another generalization will be given by the formulas:

$$\mathbf{p}^m\, \mathbf{q}^n = \sum_{j=0}^{m,n} j!\binom{m}{j}\binom{n}{j}\left(\frac{h}{2\pi i}\right)^j \mathbf{q}^{n-j}\, \mathbf{p}^{m-j},$$

$$\mathbf{q}^n\, \mathbf{p}^m = \sum_{j=0}^{m,n} j!\binom{m}{j}\binom{n}{j}\left(\frac{-h}{2\pi i}\right)^j \mathbf{p}^{m-j}\, \mathbf{q}^{n-j},$$

in which *j* increases to the smaller of the numbers *m*, *n*.

$$\mathbf{p}^m \mathbf{q}^n - \mathbf{p}^n \mathbf{q}^m = n\frac{h}{2\pi i}\sum_{l=0}^{n-1} \mathbf{p}^{m-1-j}\,\mathbf{q}^{n-1}\,\mathbf{p}^j \ . \qquad (48')$$

A comparison with (48) will yield:

$$\frac{1}{s+1}\sum_{l=0}^{s} \mathbf{p}^{s-l}\,\mathbf{q}^r\,\mathbf{p}^l = \frac{1}{r+1}\sum_{j=0}^{r} \mathbf{q}^{r-j}\,\mathbf{p}^s\,\mathbf{q}^j \ . \qquad (49)$$

We now assert that from (47), $\mathbf{H}^*$ is associated with:

$$\mathbf{H} = \frac{1}{s+1}\sum_{l=0}^{s} \mathbf{p}^{s-l}\,\mathbf{q}^r\,\mathbf{p}^l \ . \qquad (50)$$

We must prove only (40), in which we must recall formula (18′) from § **2**.
Now, from (50), one will have:

$$\mathbf{H}\,\mathbf{p} - \mathbf{p}\,\mathbf{H} = \frac{1}{s+1}(\mathbf{q}^r\,\mathbf{p}^{s+1} - \mathbf{p}^{s+1}\,\mathbf{q}^r),$$

and from (48) that is equivalent to the lower equation in (40).
With the use of (49), we will further get:

$$\mathbf{H}\,\mathbf{q} - \mathbf{q}\,\mathbf{H} = \frac{1}{r+1}(\mathbf{p}^s\,\mathbf{q}^{r+1} - \mathbf{q}^{r+1}\,\mathbf{p}^s),$$

which, from (48′), is equivalent to the upper equation (40). The desired proof is complete with that.

Whereas in classical mechanics, the constancy of energy $\dot{H} = 0$ can be read off immediately from the canonical equations, the law of energy $\dot{\mathbf{H}} = 0$ in quantum mechanics is much less transparent, as one sees.

One will recognize the extent to which its provability is far from trivial, given the assumptions that were made, when one seeks to prove the constancy of $\mathbf{H}$ simply by calculating $\dot{\mathbf{H}}$, which is closer to the classical method of proof. To that end, one must first represent $\dot{\mathbf{H}}$ as a function of $\mathbf{p}$, $\mathbf{q}$ and $\dot{\mathbf{p}}$, $\dot{\mathbf{q}}$ by means of (11), (11′), into which one must introduce the values $\frac{\partial \mathbf{H}}{\partial \mathbf{q}}$, $\frac{\partial \mathbf{H}}{\partial \mathbf{p}}$ for $\dot{\mathbf{p}}$, $\dot{\mathbf{q}}$, resp. That will yield $\dot{\mathbf{H}}$ as a function of $\mathbf{p}$ and $\mathbf{q}$. Equation (38) (the one that is derived from it in the formulas that are cited in the footnote on the previous page, resp.) allows one to recalculate this function as a sum of terms $a\,\mathbf{p}^s\,\mathbf{q}^r$ and to prove that the coefficient a of each such term will vanish. This calculation will be developed ([1]) for the most general case that was treated above in

([1]) This can be accomplished with the help of (39′) for the case of $\mathbf{H} = \frac{1}{2m}\mathbf{p}^2 + \mathbf{U}(\mathbf{q})$.

another way everywhere that it hardly seems to be practicable. Nonetheless, if the laws of energy and frequency can be proved in such a general context, then it would seem to us that this would give one a strong support for the hope that this theory actually includes deep physical laws.

In conclusion, let us note just one result that is easy to infer from the formulas of this paragraph:

Equations (35), (37) *can be replaced with* (38) *and* (44) (in which **H** means *energy*); *the frequencies are then to be determined from the frequency condition.*

We shall go into the important applications that this theorem admits in the continuation to this paper.

Chapter III – Examination of the anharmonic oscillator

Heisenberg has already considered the anharmonic oscillator with:

$$\mathbf{H} = \frac{1}{2}\mathbf{p}^2 + \frac{\omega_0^2}{2}\mathbf{q}^2 + \frac{1}{3}\lambda\mathbf{q}^3 \tag{51}$$

thoroughly. Nevertheless, we shall devote a new investigation to it here, and indeed with the goal of establishing the *most general* solution of the basic equations for that case. If the basic equations of the theory were actually complete and required no further extension then the absolute values $| q(nm) |$, $| p(nm) |$ of the components of **q** and **p**, resp., would have to be established *uniquely* by them, and it would be important to test this with the example (51). By contrast, we should expect that an indeterminacy would still exist in regard to the *phases* φ_{nm}, ψ_{nm} in:

$$q(n\,m) = | q(n\,m) | \, e^{i\varphi_{nm}},$$
$$p(n\,m) = | p(n\,m) | \, e^{i\psi_{nm}}.$$

For statistics – e.g., the interaction of quantum atoms with external radiation fields – it would be of fundamental significance to establish the degree of this indeterminacy precisely.

§ 5. Harmonic oscillator. – The starting point of our argument is the theory of the harmonic oscillator. For small λ, one can regard the motion according to equation (51) as a perturbation of the harmonic oscillation with the energy:

$$\mathbf{H} = \frac{1}{2}\mathbf{p}^2 + \frac{\omega_0^2}{2}\mathbf{q}^2. \tag{52}$$

An extension of **Heisenberg**'s consideration is necessary for even this simple problem. An essential statement on the form of the solution can be inferred from an analogous argument. Namely, since classically only *one* harmonic component is present, **Heisenberg** assumed that the matrix represented only the transition between neighboring states, so it would have the form:

$$\mathbf{q} = \begin{pmatrix} 0 & q^{(01)} & 0 & 0 & 0 & \cdots \\ q^{(10)} & 0 & q^{(12)} & 0 & 0 & \cdots \\ 0 & q^{(21)} & 0 & q^{(23)} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{53}$$

We endeavor to construct the entire theory autonomously, without appealing to the classical theory on the grounds of the correspondence principle. We will then investigate whether or not the form (53) of the matrix can be derived from the basic equations themselves, or which additional demands must be imposed when that is not the case, resp.

One sees, with no further discussion, from what was said in § **3** on invariance under permutations of rows and columns that the exact form of the matrix (53) can never be deduced from the basic equations. If one then switches the rows and columns in the same way then the canonical equations and the quantum condition will remain invariant, so one would have then found a new, apparently distinct, solution in that way. We would like to prove that the solution can always be brought into the form (53) by a mere renumbering of the elements. The equation of motion:

$$\ddot{\mathbf{q}} + \omega_0^2 \, \mathbf{q} = 0 \tag{54}$$

reads:

$$(\nu^2 (n\,m) - \nu_0^2)\, q\,(n\,m) = 0 \tag{55}$$

for the elements, in which:

$$\omega_0 = 2\pi\,\nu_0\,, \qquad h\nu\,(n\,m) = W_n - W_m\,.$$

It will follow from the sharpened quantum condition:

$$\mathbf{p}\,\mathbf{q} - \mathbf{q}\,\mathbf{p} = \frac{h}{2\pi i}\,\mathbf{1} \tag{56}$$

that for each n there must exist an n' such that $q\,(n\,n') \neq 0$. If there were an n for which all $q\,(n\,n')$ would be zero then the n^{th} diagonal term of $\mathbf{p}\,\mathbf{q} - \mathbf{q}\,\mathbf{p}$ would be equal to zero, which would contradict the quantum condition. (55) then implies that an n' will always exist for which one has:

$$|\,W_n - W_{n'}| = h\nu_0\,.$$

However, since we have assumed in our basic principles that one always has $W_n \neq W_m$ for $n \neq m$, at most *two* such indices n and n' can exist. The associated $W_{n'}$, $W_{n''}$ will then be solutions of the quadratic equation:

$$(W_n - x)^2 = h^2 \nu_0^2 .$$

If *two* such indices n', n'' actually exist then it will follow that the associated frequencies will obey:

$$\nu(n\,n') = -\,\nu(n\,n''). \tag{57}$$

Moreover, from (56), one will have:

$$\sum_k \nu(kn)\,|\,q(nk)\,|^2 = \nu(n'n)\,\{|\,q(n\,n')\,|^2 - |\,q(n\,n'')\,|^2\} = \frac{h}{8\pi^2}, \tag{58}$$

and the energy (52) will become:

$$H(n\,m) = \tfrac{1}{2} 4\pi^2 \sum_k \{-\nu(n\,n')\;\nu(k\,m)\;q(n\,k)\;q(k\,m) + \nu_0^2\, q(n\,k)\;q(k\,m)\}$$

$$= 2\pi^2 \sum_k q(n\,k)\,q(k\,m)\{\nu_0^2 - \nu(n\,k)\;\nu(k\,m)\}.$$

In particular, for $m = n$, one will have:

$$H(n\,n) = W_n = 4\pi^2 \nu_0^2 (|\,q(n\,n')\,|^2 + |\,q(n\,n'')\,|^2). \tag{59}$$

Now, three further cases are possible:

a) There is no n'', and one has $W_{n'} > W_n$.

b) There is no n'', and one has $W_{n'} < W_n$.

c) There exists an n''.

In case *b*), we now consider n', instead of n. At most two indices $(n')'$ and $(n')''$ belong to it, and one of them must be equal to n. With that, we come back to one of the cases *a*) or *c*), and for that reason, we can ignore *b*).

In case *a*), one will have $\nu(n'n) = +\,\nu_0$, and it will follow from (58) that:

$$\nu_0 \cdot |\,q(n\,n')\,|^2 = \frac{h}{8\pi^2}, \tag{60}$$

so from (59):

$$W_n = H(n\,n) = 4\pi^2 \nu_0^2\,|\,q(n\,n')\,|^2 = \tfrac{1}{2}\,\nu_0\,h.$$

Due to the assumption $W_n \neq W_m$ for $n \neq m$, there is then at most *one* index $n = n_0$ for which case *a*) is true.

If such an n_0 exists then we can give a sequence of numbers:

$$n_0, n_1, n_2, n_3, \ldots$$

in such a way that:

$$(n_k)' = n_{k+1} \qquad \text{and} \qquad W_{k+1} > W_k .$$

One then always have:

$$(n_{k+1})'' = n_k .$$

Hence, for $k > 0$, (58) and (59) will imply that:

$$H(n_k\, n_k) = 4\pi^2 \nu_0^2 \{ |\, q(n_k, n_{k+1})\, |^2 + |\, q(n_k, n_{k-1})\, |^2 \}, \tag{61}$$

$$\tfrac{1}{2} h = 4\pi^2 \nu_0 \{ |\, q(n_k, n_{k+1})\, |^2 - |\, q(n_k, n_{k-1})\, |^2 \}. \tag{62}$$

It follows from (60) and (62) that:

$$|\, q(n_k, n_{k+1})\, |^2 = \frac{h}{8\pi^2 \nu_0} (k+1), \tag{63}$$

and it will then follow from (61) that:

$$W_{n_k} = H(n_k, n_k) = \nu_0 h (k + \tfrac{1}{2}). \tag{64}$$

Now, we would still like to see whether it is possible that there is no n for which case *a*) is valid. Beginning with an arbitrary n_0, we can then define $n_0' = n_1$ and $n_0'' = n_{-1}$; for each of them, we again define $n_1' = n_2$, $n_1'' = n_0$, and $n_{-1}' = n_0$, $n_{-1}'' = n_{-2}$, etc. In that way, we will get a sequence of numbers:

$$\ldots, n_{-2}, n_{-1}, n_0, n_1, n_2, \ldots, \tag{65}$$

and equations (61), (62) will be true for every k between $-\infty$ and $+\infty$. However, that is impossible. From (62), the quantities $x_k = |\, q(n_{k+1}, n_k)\, |^2$ will then define an equidistant number sequence, and since they are positive, it must have a smallest one. We can once more denote the corresponding index by n_0 and then come back to the previous case; formulas (63), (64) are also valid here.

One further sees that every number n must be included in the numbers n_k; otherwise, one could define a new sequence (65) with n as its starting term, which would again make formula (60) true. The starting terms of both sequences would then have the same value $W_n = H(n\, n)$, which is impossible.

With that, we have proved that the indices 0, 1, 2, 3, … can be reordered into a new sequence $n_0, n_1, n_2, n_3, \ldots$ such that the solution has the **Heisenberg** form (53). That would then seem to be the "normal form" for the general solution. From (64), it has the property that:

$$W_{n_{k+1}} > W_{n_k}\,.$$

Conversely, if one demands that $W_n = H\,(n\,n)$ should always increase with n that one will necessarily have $n_k = k$; this principle then establishes the normal form uniquely. However, only the notation will be fixed in that way, and the calculation will take a more transparent form; *physically*, nothing new will come about in that way.

In this, we have a deeper difference from the semi-classical determination of the stationary states that has been used up to now. The classically-calculated paths close into each other continuously, whereby the quantum paths that were subsequently rejected will also define a well-defined sequence from the outset. The new mechanics represents a true theory of the discontinuum, in which no mention is made of the sequence of quantum states that are defined by the physical process, as such, since the quantum numbers are really nothing but distinctive indices that one can order and normalize according to whatever practical viewpoint one desires (e.g., in increasing energy W_n).

§ 6. Anharmonic oscillator. – The equations of motion:

$$\ddot{\mathbf{q}} + \omega_0^2\,\mathbf{q} + \lambda\,\mathbf{q}^2 = 0, \tag{66}$$

together with the quantum condition, give the following system of equations for the elements:

$$\left.\begin{aligned} (\omega_0^2 - \omega^2(nm))\,q(nm) + \lambda\sum_k q(nk)\,q(km) = 0, \\ \sum_k \omega(nk)\,q(nk)\,q(kn) = -\frac{h}{4\pi}. \end{aligned}\right\} \tag{67}$$

We shall attempt to solve this by a series development:

$$\left.\begin{aligned} \omega(n\,m) = \omega^0(n\,m) + \lambda\,\omega^{(1)}(n\,m) + \lambda^2\,\omega^{(2)}(n\,m) + \cdots, \\ q(n\,m) = q^0(n\,m) + \lambda\,q^{(1)}(n\,m) + \lambda^2\,q^{(2)}(n\,m) + \cdots \end{aligned}\right\} \tag{68}$$

For $\lambda = 0$, one has the case of the harmonic oscillator that was treated in the previous paragraph. We write the solution (53) in the form:

$$q^0\,(n\,m) = a_n\,\delta_{n,\,m-1} + \overline{a_m}\,\delta_{n-1,\,m}\,, \tag{69}$$

in which the overbar denotes the complex-conjugate quantity. If one takes the square and higher powers of the matrix $\mathbf{q}^0 = (q^0\,(n\,m))$ then matrices of a similar form will appear, namely, sums of terms:

$$(\xi)^{(p)}_{nm} = \xi_n\,\delta_{n,\,m-p} + \overline{\xi_m}\,\delta_{n-p,\,m}\,. \tag{70}$$

One will then be in a good position to put the solution into the form:

$$\left.\begin{aligned} q^0(n\,m) &= (a)^{(1)}_{nm}, \\ q^{(1)}(n\,m) &= (x)^0_{nm} + (x')^{(2)}_{nm}, \\ q^{(2)}(n\,m) &= (y)^{(1)}_{nm} + (y')^{(3)}_{nm}, \\ &\cdots\cdots\cdots\cdots \end{aligned}\right\} \tag{71}$$

in which one always alternates between odd and even values of the index p.

In fact, if one inserts this into the approximate equations:

$$\lambda:\left\{\begin{aligned} &(\omega_0^2 - \omega^0(n\,m)^2)\,q^{(1)}(n\,m) - 2\omega^0(n\,m)\,\omega^{(1)}(n\,m)\,q^0(n\,m) + \sum_k q^0(n\,k)\,q^0(k\,m) = 0, \\ &\sum_k \{\omega^0(n\,k)\,(q^0(n\,k)\,q^{(1)}(k\,n) + q^{(1)}(n\,k)\;q^0(k\,n) + \omega^{(1)}(n\,k)\,q^0(n\,k)\,q^0(k\,n)\} = 0, \end{aligned}\right\} \tag{72}$$

$$\lambda^2:\left\{\begin{aligned} &(\omega_0^2 - \omega^0(n\,m)^2)\,q^{(2)}(n\,m) - 2\omega^0(n\,m)\,\omega^{(1)}(n\,m)\,q^{(1)}(n\,m) - (\omega^0(n\,m)^2 \\ &\quad + 2\omega^0(n\,m)\,\omega^{(2)}(n\,m)\,q^0(n\,m) + \sum_k\; (q^0(n\,k)\,q^{(1)}(k\,m) + q^{(1)}(n\,k)\;q^0(k\,m)) = 0 \\ &\sum_k \{\omega^0(n\,k)\,(q^0(n\,k)\,q^{(2)}(k\,m) + q^{(1)}(n\,k)\;q^{(1)}(k\,m) + q^{(2)}(n\,k)\,q^0(k\,n)) \\ &\quad + \omega^{(1)}(n\,k)\,(q^0(n\,k)\,q^{(1)}(k\,m) + q^{(1)}(n\,k)\,q^0(k\,m)) + \omega^{(2)}(n\,k)\,q^0(n\,k)\,q^0(k\,m) = 0 \end{aligned}\right\} \tag{73}$$

and observes the multiplication rule:

$$\left.\begin{aligned} \sum_k \Omega_{nkm}\,(\xi)^{(p)}_{nk}\,(\eta)^{(q)}_{km} &= \Omega_{n,n+p,n+p+q}\,\xi_n\,\eta_{m+p}\,\delta_{n,m-p-q} \\ &+ \Omega_{n,n+p,n+p-q}\,\xi_n\,\overline{\eta}_{n+p-q}\,\delta_{n,m-p+q} \\ &+ \Omega_{n,n-p,n-p+q}\,\overline{\xi}_{n-p}\,\eta_{n-p}\,\delta_{n,m+p-q} \\ &+ \Omega_{n,n-p,n-p-q}\,\overline{\xi}_{n-p}\,\eta_{n-p-q}\,\delta_{n,m+p+q} \end{aligned}\right\} \tag{74}$$

then when one sets the individual factors of $\delta_{n,\,m-s}$ equal to zero, one will see that all conditions can be fulfilled exactly by the Ansatz (71), and the higher terms in (71) will vanish identically.

In the individual cases, calculation will give the following:

After one substitutes the expressions in (71), the first of equations (72) will yield:

$$\left.\begin{aligned} 2\omega_0^2\,x_n + |\,a_n\,|^2 + |\,a_{n-1}\,|^2 &= 0, \\ -3\,\omega_0^2\,x_{n'} + a_n a_{n+1} &= 0, \\ \omega^{(1)}_{n,n-1} &= 0, \end{aligned}\right\} \tag{75}$$

while the second one will be fulfilled identically. One will then have:

$$\left.\begin{aligned} x_n &= -\frac{|a_n|^2 + |a_{n-1}|^2}{2\omega_0^2}, \\ x_n' &= \frac{a_n a_{n+1}}{3\omega_0^2}. \end{aligned}\right\} \tag{76}$$

The first of equations (73) yields:

$$\left.\begin{aligned} 2\omega_0 a_n \omega_{n,n+1}^{(2)} + 2a_n x_{n+1} + 2a_n x_n + \bar{a}_{n-1} x_{n-1}' + a_{n+1} x_n' &= 0, \\ -8\omega_0^2 y_n' + a_n x_{n+1}' + a_{n+2} x_n' &= 0, \\ \omega_{n,n-2}^{(1)} &= 0, \end{aligned}\right\} \tag{77}$$

while the second equation is not fulfilled identically, but implies a determining equation for y_n :

$$\left.\begin{aligned} a_n \bar{y}_n + \bar{a}_n y_n - a_{n-1} \bar{y}_{n-1} - \bar{a}_{n-1} y_{n-1} + 2|x'|^2 - 2|x_{n-2}'|^2 \\ - \frac{\omega_{n,n+1}^{(2)}}{\omega_0} |a_n|^2 - \frac{\omega_{n,n-1}^{(2)}}{\omega_0} |a_{n-1}|^2 = 0. \end{aligned}\right\} \tag{78}$$

The solution reads:

$$\left.\begin{aligned} \omega_{n,n+1}^{(2)} &= \frac{1}{3\omega_0^3}(|a_{n+1}|^2 + |a_{n-1}|^2 + 3|a_n|^2), \\ y_n' &= \frac{1}{12\omega_0^4} a_n a_n a_{n+2}. \end{aligned}\right\} \tag{79}$$

If one further sets:

$$\eta_n = a_n \bar{y}_n + \bar{a}_n y_n, \tag{80}$$

to abbreviate, then the η can be determined from the equation:

$$\eta_n - \eta_{n-1} = \frac{1}{\omega_0^4}(|a_n|^4 - |a_{n-1}|^4 + \frac{1}{9}|a_n|^2 |a_{n+1}|^2 - \frac{1}{9}|a_{n-1}|^2 |a_{n-2}|^2). \tag{81}$$

The expressions (76) and (79) show that the quantities x_n, x_n', y_n' can be expressed by the solution of the zeroth approximation a_n . Their phases can then be established by those of the harmonic oscillator. Things seem to be otherwise for the quantities y_n ; indeed, η_n is then determined uniquely by (81), but y_n cannot be established completely from (80). It is likely that an extended determining equation for y_n will arise from the following approximations. We must leave that question open here, but we would like to mention its essential meaning for the sake of the unity of the entire theory. It will then happen that no statistical questions at all will depend upon whether our conjecture is correct that of the phases of the q (n m), *one* of them in each row (or column) of the matrix remain undetermined.

In conclusion, we would like to give the explicit formulas that one obtains when one substitutes the previously-found (§ **5**) solution for the harmonic oscillator. From (63), they read:

$$a_n = \sqrt{C\,(n+1)}\; e^{i\varphi_n}, \qquad C = \frac{h}{4\pi\omega_0} = \frac{h}{8\pi^2\nu_0} \tag{82}$$

in normal form. With that, one gets from (76), (79), (81):

$$\left.\begin{aligned} x_n &= -\frac{C}{2\omega_0^2}(2n+1), \\ x'_n &= \frac{C}{3\omega_0^2}\sqrt{(n+1)(n+2)}\; e^{i(\varphi_n+\varphi_{n+1})}, \\ y'_n &= \frac{\sqrt{C^3}}{12\omega_0^4}\sqrt{(n+1)(n+2)(n+3)}\; e^{i(\varphi_n+\varphi_{n+1}+\varphi_{n+2})}, \end{aligned}\right\} \tag{83}$$

$$\left.\begin{aligned} \omega^{(1)}_{n,n-1} = 0, \quad \omega^{(1)}_{n,n-2} = 0, \\ \omega^{(2)}_{n,n-1} = -\frac{5}{3}\frac{C}{\omega_0^3}n; \end{aligned}\right\} \tag{84}$$

hence:

$$\eta_n - \eta_{n-1} = \frac{11}{9}\frac{C^2}{\omega_0^4}(2n+1),$$

$$\eta_n = a_n\,\bar{y}_n + \bar{a}_n\,y_n = \frac{11}{9}\frac{C^2}{\omega_0^4}(n+1)^2.$$

If one sets $y_n = |\,y_n\,|\; e^{i\varphi_n}$ then one will have:

$$|\,y_n\,|\cos(\varphi_n - \psi_n) = \frac{\eta_n}{2\,|\,a_n\,|} = \frac{11}{18}\frac{\sqrt{C}^3}{\omega_0^4}\sqrt{n+1}^3. \tag{85}$$

Nothing more can be said about y_n in this approximation.

However, we would like to write out the final formulas under the assumption that $\psi_n = \varphi_n$. They then read (up to terms of order higher than two in λ):

$$\left.\begin{aligned} \omega(n, n-1) &= \omega_0 - \lambda^2\frac{5}{3}\frac{C}{\omega_0^3}n + \cdots, \\ \omega(n, n-2) &= 2\omega_0 + \cdots; \end{aligned}\right\} \tag{86}$$

$$\left.\begin{aligned} q(n,n) &= -\lambda\frac{C}{\omega_0^2}(2n+1)+\cdots, \\ q(n,n-1) &= \sqrt{C\,n}\,e^{i\varphi_{n-1}}\left(1+\lambda^2\frac{11}{18}\frac{Cn}{\omega_0^4}+\cdots\right), \\ q(n,n-2) &= \lambda\frac{C}{\omega_0^2}\sqrt{n(n-1)}\;e^{i(\varphi_{n-1}+\varphi_{n-2})}+\cdots, \\ q(n,n-3) &= \lambda^2\frac{\sqrt{C}^3}{12\omega_0^4}\sqrt{n(n-1)(n-2)}\;e^{i(\varphi_{n-1}+\varphi_{n-2}+\varphi_{n-3})}+\cdots \end{aligned}\right\} \qquad (87)$$

We have also calculated the energy directly and found that:

$$W_n = h\nu_0\left(n+\frac{1}{2}\right)-\lambda^2\frac{5}{3}\frac{C^2}{\omega_0^2}\left(n(n+1)+\frac{17}{30}\right)+\ldots \qquad (88)$$

The frequency condition is, in fact, fulfilled, since one has, when one goes back to (82):

$$W_n - W_{n-1} = h\nu_0 - \lambda^2\frac{2C^2}{\omega_0^2}n + \ldots = \frac{h}{2\pi}\,\omega(n,\,n-1),$$

$$W_n - W_{n-2} = 2h\nu_0 + \ldots \qquad = \frac{h}{2\pi}\,\omega(n,\,n-2).$$

One can, with **Heisenberg**, connect formula (88) with the remark that a deviation from the classical theory exists already in the terms of lowest order that one can rectify formally with a "half-integer" quantum number $n' = n + 1/2$. Moreover, our expressions $\omega(n,\,n-1)$ agree with (86) and the classical frequencies *precisely*. The classical energy is then ([1]):

$$W_n^{(kl)} = h\nu_0\,n - \lambda^2\cdot\frac{5}{3}\frac{C^2}{\omega_0^2}n^2 + \ldots,$$

and thus the classical frequencies will be:

$$\omega_{kl} = \frac{1}{h}\frac{\partial W_n^{(kl)}}{\partial n} = h\nu_0\,n - \lambda^2\cdot\frac{5}{3}\frac{C^2}{\omega_0^2}n + \ldots$$

$$= \omega_{qu}\,(n,\,n-1) = \frac{1}{h}(W_n^{(qu)} - W_{n-1}^{(qu)})\,.$$

([1]) See **M. Born**, *Atommechanik*, Berlin, 1925, Chapter 4, § **42**, pp. 294. One must set $a = 1/3$ in formula (6) in order to come into agreement with our Ansatz.

We have ultimately shown that the expression (88) can also be obtained from the **Kramers-Born** perturbation formula (up to the additive constant).

Chapter IV – Remarks on electrodynamics

According to **Heisenberg**, the square of the absolute value $|\, q\,(n\,m)\,|^2$ of the element of **q** for the case in which **q** is a Cartesian coordinate is definitive of the *jump probabilities.* Here, to conclude, we would like to show the way by which this assumption can be obtained as a foundation from more general arguments. In order to do that, it will be necessary to address the question of how the basic equations of electrodynamics are to be reinterpreted in the sense of the new theory. However, we would like to emphasize that the argument that will be presented here has only a heuristic character; they only serve to help us recognize our fundamental viewpoint on the problem. A thorough treatment of the questions that appear here shall be given later, in which, above all, the relationship of the theory that is put forth to the theory of light quanta will be discussed.

Here, we would like to discuss only those points that can be reached without going into the exact form of the quantum condition for systems of many degrees of freedom. One can see that one can already go rather far into electrodynamics in that way by the following argument: The electromagnetically-oscillating *cavity* represents a system with *infinitely many degrees of freedom.* Nonetheless, the basic theorems that were developed in the foregoing chapters, which indeed refer to only systems of *one* degree of freedom, are sufficient for its treatment, because it will go to a system of *uncoupled* oscillators after it is analyzed into eigen-oscillations. It is hardly possible to doubt the way that this system should be treated. For it, the fact that the basic electromagnetic equations are linear (viz., the principle of superposition) proves to have special significance. It will then follow from that fact that the virtual oscillators are *harmonic*, and the validity of the law of energy will be independent of the quantum condition for harmonic oscillators precisely (as opposed to the behavior of other systems). It follows from:

$$\mathbf{H} = \tfrac{1}{2}(\mathbf{p}^2 + \omega_0^2\,\mathbf{q}^2)$$

that

$$\begin{aligned}\dot{\mathbf{H}} &= \tfrac{1}{2}(\dot{\mathbf{p}}\mathbf{p} + \mathbf{p}\dot{\mathbf{p}} + \omega_0^2\,\dot{\mathbf{q}}\mathbf{q} + \omega_0^2\,\mathbf{q}\dot{\mathbf{q}}) \\ &= \tfrac{1}{2}\omega_0^2(-\mathbf{q}\mathbf{p} - \mathbf{p}\mathbf{q} + \mathbf{p}\mathbf{q} + \mathbf{q}\mathbf{p}) \\ &= 0.\end{aligned}$$

Correspondingly, one would then expect that the integral laws of the electrodynamics of the vacuum (viz., the laws of energy and impulse) could be arrived at in a completely general way from the **Maxwell** equations alone, when reinterpreted in terms of matrices, without having to go into the quantum condition. When we have shown that, we will likewise obtain the means to give a basis for **Heisenberg**'s statement about the meaning of $|\, q\,(n\,m)\,|^2$

§ 7. Maxwell's equations. Law of energy and impulse. – We would like to agree in advance that *vectors*, as usual, will always be denoted by German symbols, while the difference between numbers and matrices will still be indicated by ordinary and boldface characters. We choose our units of measurements in conjunction with **Abraham**'s textbook ([1]).

The electromagnetic processes *in vacuo* will be represented as the superposition of plane waves. For such plane waves, we will regard the electric and magnetic field strengths $\mathfrak{E}$, $\mathfrak{H}$ as *matrices* whose elements are harmonically-oscillating plane waves, so, e.g., for a suitable position of the coordinate system:

$$\mathfrak{E} = \left(\mathfrak{E}(nm)\, e^{2\pi i \nu(nm)(t-x/c)}\right). \tag{89}$$

Of course, one must realize that n, m are generally no longer restricted to a discrete set of values in this, and that they also no longer refer to single numbers, but systems of numbers (viz., vectors).

One preserves the **Maxwell** equations in the form of matrix equations:

$$\operatorname{rot} \mathfrak{H} - \frac{1}{c}\dot{\mathfrak{E}} = 0, \qquad \operatorname{rot} \mathfrak{E} + \frac{1}{c}\dot{\mathfrak{H}} = 0. \tag{90}$$

The differentiations with respect to x, y, z, t are then to be thought of as being performed on each individual element of the matrix ([2]).

We would now like to derive the law of energy-impulse. In order to do that, it will be necessary to make some prefatory remarks on the multiplication of matrix vectors.

We define the *scalar product* by:

$$(\mathfrak{A}, \mathfrak{B}) = \mathfrak{A}\mathfrak{B} = \mathfrak{A}_x \mathfrak{B}_x + \mathfrak{A}_y \mathfrak{B}_y + \mathfrak{A}_z \mathfrak{B}_z \tag{91}$$

and the *vector product* by:

$$[\mathfrak{A}\, \mathfrak{B}]_x = \mathfrak{A}_y \mathfrak{B}_z - \mathfrak{A}_z \mathfrak{B}_y\,. \tag{92}$$

Since matrix multiplication is not commutative, the relations:

$$\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A}, \qquad [\mathfrak{A}\mathfrak{B}] = -[\mathfrak{B}\mathfrak{A}]$$

are *not* generally true.

By contrast, we assert that:

$$\operatorname{div}[\mathfrak{A}\mathfrak{B}] = (\operatorname{rot}\mathfrak{A}, \mathfrak{B}) - (\mathfrak{A}, \operatorname{rot}\mathfrak{B}). \tag{93}$$

([1]) **M. Abraham**, *Theorie der Elektrizität*, II, Leipzig, 1914.

([2]) In some situations, another way of conceptualizing the electromagnetic field is required, for which the spatial coordinates do not appear as numbers, but once more as matrices. That will have a corresponding alteration of the meaning of the spatial differential quotients in the **Maxwell** equations as a consequence. We shall come back to this in the continuation of this paper.

We now define the energy density **W** (as a scalar matrix) by:

$$\mathbf{W} = \frac{1}{8\pi}(\mathfrak{E}^2 + \mathfrak{H}^2). \tag{94}$$

From (11), one will then have:

$$8\pi\dot{\mathbf{W}} = \mathfrak{E}\dot{\mathfrak{E}} + \dot{\mathfrak{E}}\mathfrak{E} + \mathfrak{H}\dot{\mathfrak{H}} + \dot{\mathfrak{H}}\mathfrak{H},$$

and from (90):

$$\frac{8\pi}{c}\mathbf{W} = (\mathfrak{E}, \text{rot } \mathfrak{H}) + (\text{rot } \mathfrak{H}, \mathfrak{E}) - (\mathfrak{H}, \text{rot } \mathfrak{E}) - (\text{rot } \mathfrak{E}, \mathfrak{H}),$$

so, from (93):

$$\dot{\mathbf{W}} + \text{div } \mathfrak{S} = 0, \tag{95}$$

in which:

$$\mathfrak{S} = \frac{c}{8\pi}([\mathfrak{E}\ \mathfrak{H}] - [\mathfrak{H}\ \mathfrak{E}]). \tag{96}$$

This is *Poynting's law* for matrix electrodynamics; $\mathfrak{S}$ means the *radiation vector*.

The law of impulse can be derived similarly: One defines the *Maxwell stresses by:*

$$\left.\begin{aligned}\mathbf{T}_{xz} &= \frac{1}{8\pi}(\mathfrak{E}_x^2 - \mathfrak{E}_y^2 - \mathfrak{E}_z^2) + (\mathfrak{H}_x^2 - \mathfrak{H}_y^2 - \mathfrak{H}_z^2),\\ \mathbf{T}_{yz} &= \frac{1}{8\pi}(\mathfrak{E}_y\mathfrak{E}_z + \mathfrak{E}_z\mathfrak{E}_y + \mathfrak{H}_y\mathfrak{H}_z + \mathfrak{H}_z\mathfrak{H}_y),\end{aligned}\right\} \tag{97}$$

and the *impulse density* of the radiation by:

$$\mathfrak{g} = \frac{1}{c^2}\mathfrak{S} = \frac{1}{8\pi c}([\mathfrak{E}\ \mathfrak{H}] - [\mathfrak{H}\ \mathfrak{E}]). \tag{98}$$

One then gets:

$$\dot{\mathfrak{g}} = \frac{\partial \mathbf{T}_{xx}}{\partial x} + \frac{\partial \mathbf{T}_{xy}}{\partial y} + \frac{\partial \mathbf{T}_{xz}}{\partial z} \tag{99}$$

by a similar calculation.

Naturally, one will arrive at these relations more intuitively when one employs the four-dimensional representation of the theory of relativity. A systematic treatment of the four-dimensional vector analysis and the theory of relativity on the basis of matrix theory, with its non-commutative multiplication, shall be given at a later point.

§ 8. Spherical waves. Radiation from a dipole. – In order to pursue our goal of calculating the radiation of an oscillator, we must now draw our attention to *spherical waves.*

For that purpose, we shall introduce the *Hertzian vector* $\mathfrak{Z}$ as a matrix vector. We get $\mathfrak{E}$ and $\mathfrak{H}$ from $\mathfrak{Z}$ by means of the equations:

$$\mathfrak{E} = \text{grad div}\,\mathfrak{Z} - \frac{1}{c^2}\ddot{\mathfrak{Z}}, \qquad \mathfrak{H} = \frac{1}{c}\text{rot}\,\dot{\mathfrak{Z}}. \tag{100}$$

In the classical theory, $\mathfrak{Z}$ is proportional to:

$$\frac{1}{r}e^{2\pi i\nu(t-r/c)}$$

for a spherical wave.

Now, it is known that this expression can be written as the superposition of plane waves ([1]) on the basis of the identity:

$$\frac{e^{i\kappa r}}{r} = \frac{i\kappa}{2\pi}\int e^{i\kappa(\mathfrak{r}\mathfrak{s})}d\omega. \tag{101}$$

In this, $\mathfrak{r}$ is the numerical vector from the center of the spherical wave to the reference point, $\mathfrak{s}$ is a unit vector, and $d\omega = d\mathfrak{s}_x\, d\mathfrak{s}_y\, d\mathfrak{s}_z$. Thus, plane waves, which are represented by matrices of the form (89), can be obtained in our theory by integrating over the directions of the wave normals in the representation of a spherical wave:

$$\mathfrak{Z} = \left(e\mathfrak{q}(nm)\frac{e^{2\pi i\nu(nm)(t-r/c)}}{r} \right); \tag{102}$$

the matrix $e\,\mathfrak{q} = (e\,\mathfrak{q}\,(n\,m))$ in this represents the electric moment that the wave produces.

The calculations that lead from here to the determination of the electromagnetic field and radiation are the same as in the classical theory, since $\mathfrak{r}$, as a numerical vector, commutes with any matrix. One obtains:

$$\left.\begin{aligned} \mathfrak{H} &= -\frac{e}{c^2}\frac{1}{r^2}[\mathfrak{r}\ddot{\mathfrak{q}}], \\ \mathfrak{E} &= \frac{e}{c^2}\frac{1}{r^2}[\mathfrak{r}[\mathfrak{r}\ddot{\mathfrak{q}}]], \end{aligned}\right\} \tag{103}$$

and that will give:

$$\mathfrak{S} = \frac{e}{4\pi c^3}\frac{\mathfrak{r}}{r^2}[\mathfrak{r}\ddot{\mathfrak{q}}]. \tag{104}$$

([1]) See, e.g., **P. Debye**, Ann. Phys. (Berlin) **30** (1909), 755; Formula (7″), pp. 758.

The integration over all spatial directions proceeds in the same way as it does in the classical theory. The result for the energy that is radiated per second reads:

$$\int \mathfrak{S}\, d\mathfrak{f} = \frac{2e^2}{3c^3}\ddot{\mathbf{q}}. \tag{105}$$

In order to obtain the mean radiation, one must take the mean of this expression over time; the diagonal matrix:

$$\frac{2e^2}{3c^3}\overline{\ddot{\mathbf{q}}^2} \tag{106}$$

will then arise in that way.

If the oscillator oscillates in a fixed direction then we can replace the matrix vector $\mathbf{q}$ with the matrix scalar $\mathbf{q} = (q\,(n\,m))$; the radiation will then become:

$$\frac{2e^2}{3c^3}\overline{\ddot{\mathbf{q}}^2} = \frac{32\pi^4 e^2}{3c^3}\left(\sum_k \nu(nk)^4\,|\,q(nk)\,|^2\right). \tag{107}$$

We still cannot give a complete theory of radiation here, from which one could inevitably go from the ordering of the individual terms in this series to the stationary states. A more precise examination of the reaction of the radiation on the oscillator would be necessary for that, and therefore a theory of damping. We will come back to that later. Here, we would only like to find out whether the radiation is actually determined by the quantities $|\,q\,(n\,m)\,|^2$. The expression (107) shows that this is the case, but at the same time, we see that the quantities that are written out are not all of the radiation that emanates spontaneously from a stationary state. The spontaneous transitions always result from only lower-energy states or states with lower quantum numbers, with a suitable numbering. We can now state how this situation can be expressed in our theory in an entirely formal way: In order to do that, we do not take the mean value, but the *diagonal sum* of the radiation matrix (105); that will give:

$$D\left(\frac{2e^2}{3c^3}\ddot{\mathbf{q}}^2\right) = \frac{32\pi^4 e^2}{3c^3}\sum_{n,k}\nu\,(nk)^4\,|\,q\,(nk)\,|^2\,. \tag{108}$$

Here, we can sum the right-hand side again and write:

$$\frac{64\pi^4 e^2}{3c^3}\sum_n\left(\sum_{k<n}\nu\,(nk)^4\,|\,q\,(nk)\,|^2\right). \tag{109}$$

The desired ordering is achieved with that: Each state n belongs to the radiation that corresponds to the transitions from all states $k < n$, each of which has the intensity that is known from the classical theory. That will agree with experiment when one assumes that the indices n are ordered by increasing energies W_n .

Heisenberg's assumption, in the restricted sense that is characterized above, is then justified with that.

It should be likewise emphasized here that this convention in regard to the jump probabilities is independent of the assumption of the nondegeneracy of the system; viz., the distinctness of all W_n. In conclusion, we shall stress that the *statistical weights* of the states are established along with the transition probabilities, and indeed each state that is characterized by a row and column (a diagonal term of **W**) must be ascribed the same statistical weight. The fact that this result (when generalized to systems of many degrees of freedom) leads to the basic principle of **Bose-Einstein** statistics of light quanta by itself will be explained later.

Remark added in proof. – Meanwhile, the generalization of the theory to many degrees of freedom that was announced above has been worked out along with **W. Heisenberg** and will be presented in the continuation to this paper. Various points that were touched upon here already will be discussed more thoroughly, since they have been further clarified in the meantime.
