Received September 27, 1925

13

ON QUANTUM MECHANICS

M. BORN AND P. JORDAN

The recently published theoretical approach of Heisenberg is here developed into a systematic theory of quantum mechanics (in the first place for systems having one degree of freedom) with the aid of mathematical matrix methods. After a brief survey of the latter, the mechanical equations of motion are derived from a variational principle and it is shown that using Heisenberg's quantum condition, the principle of energy conservation and Bohr's frequency condition follow from the mechanical equations. Using the anharmonic oscillator as example, the question of uniqueness of the solution and of the significance of the phases of the partial vibrations is raised. The paper concludes with an attempt to incorporate electromagnetic field laws into the new theory.

Introduction

The theoretical approach of Heisenberg[1] recently published in this Journal, which aimed at setting up a new kinematical and mechanical formalism in conformity with the basic requirements of quantum theory, appears to us of considerable potential significance. It represents an attempt to render justice to the new facts by setting up a new and really suitable conceptual system instead of adapting the customary conceptions in a more or less artificial and forced manner. The physical reasoning which led Heisenberg to this development has been so clearly described by him that any supplementary remarks appear superfluous. But, as he himself indicates, in its formal, mathematical aspects his approach is but in its initial stages. His hypotheses have been applied only to simple examples without being fully carried through to a generalized theory. Having been in an advantageous position to familiarize ourselves with his ideas throughout their formative stages, we now strive (since his investigations have been

Editor's note. This paper was published as Zs. f. Phys. **34** (1925) 858–888. Chapter 4 (pp. 883–888) of the original paper is not reproduced here.

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

concluded) to clarify the mathematically formal content of his approach and present some of our results here. These indicate that it is in fact possible, starting with the basic premises given by Heisenberg, to build up a closed mathematical theory of quantum mechanics which displays strikingly close analogies with classical mechanics, but at the same time preserves the characteristic features of quantum phenomena.

In this we at first confine ourselves, like Heisenberg, to systems having *one degree of freedom* and assume these to be – from a classical standpoint – *periodic*. We shall in the continuation of this publication concern ourselves with the generalization of the mathematical theory to systems having an arbitrary number of degrees of freedom, as also to aperiodic motion. A noteworthy generalization of Heisenberg's approach lies in our confining ourselves neither to treatment of nonrelativistic mechanics nor to calculations involving Cartesian systems of coordinates. The only restriction which we impose upon the choice of coordinates is to base our considerations upon *libration coordinates*, which in classical theory are *periodic* functions of time. Admittedly, in some instances it might be more reasonable to employ other coordinates: for example, in the case of a rotating body to introduce the angle of rotation φ, which becomes a linear function of time. Heisenberg also proceeded thus in his treatment of the rotator; however, it remains undecided whether the approach applied there can be justified from the standpoint of a consistent quantum mechanics.

The mathematical basis of Heisenberg's treatment is the *law of multiplication* of quantum-theoretical quantities, which he derived from an ingenious consideration of correspondence arguments. The development of his formalism, which we give here, is based upon the fact that this rule of multiplication is none other than the well-known mathematical rule of *matrix multiplication*. The infinite square array (with discrete or continuous indices) which appears at the start of the next section, termed a *matrix*, is a representation of a physical quantity which is given in classical theory as a function of time. The mathematical method of treatment inherent in the new quantum mechanics is thereby characterized through the employment of *matrix analysis* in place of the usual number analysis.

Using this method, we have attempted to tackle some of the simplest problems in mechanics and electrodynamics. A *variational*

principle, derived from correspondence considerations, yields *equations of motion* for the most general Hamilton function which are in closest analogy with the classical canonical equations. The quantum condition conjoined with one of the relations which proceed from the equations of motion permits a simple matrix notation. With the aid of this, one can prove the general validity of the *law of conservation of energy* and the *Bohr frequency relation* in the sense conjectured by Heisenberg: this proof could not be carried through in its entirety by him even for the simple examples which he considered. We shall later return in more detail to one of these examples in order to derive a basis for consideration of the part played by the phases of the partial vibrations in the new theory. We show finally that the basic laws of the electromagnetic field in a vacuum can readily be incorporated and we furnish substantiation for the assumption made by Heisenberg that the squares of the absolute values of the elements in a matrix representing the electrical moment of an atom provide a measure for the transition probabilities.

CHAPTER 1. MATRIX CALCULATION

1. Elementary operations. Functions

We consider square infinite matrices,[1] which we shall denote by heavy type to distinguish them from ordinary quantities which will throughout be in light type,

$$\mathbf{a} = (a(nm)) = \begin{pmatrix} a(00) & a(01) & a(02) \dots \\ a(10) & a(11) & a(12) \dots \\ a(20) & a(21) & a(22) \dots \\ \dots & \dots & \dots \end{pmatrix}.$$

Equality of two matrices is defined as equality of corresponding components:

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

Matrix addition is defined as addition of corresponding components:

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

[1] Further details of matrix algebra can be found, e.g., in M. Bôcher, Einführung in die höhere Algebra (translated from the English by Hans Beck; Teubner, Leipzig, 1910) § 22–25; also in R. Courant and D. Hilbert, Methoden der mathematischen Physik **1** (Springer, Berlin, 1924) Chapter I.

Matrix multiplication is defined by the rule 'rows times columns', familiar from the theory of determinants:

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

Powers are defined by repeated multiplication. The associative rule applies to multiplication and the distributive rule to combined addition and multiplication:

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

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

However, the commutative rule does *not* hold for multiplication: it is not in general correct to set $ab=ba$. If a and b do satisfy this relation, they are said to commute.
The *unit matrix* defined by

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

has the property

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

The *reciprocal matrix* to a, namely a^{-1}, is defined by[1]

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

As *mean value* of a matrix a we denote that matrix whose diagonal elements are the same as those of a whereas all other elements vanish:

$$\bar{a} = (\delta_{nm} a(nn)). \tag{8}$$

The sum of these diagonal elements will be termed the *diagonal sum* of the matrix a and written as $D(a)$, viz.

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

From (3) it is easy to prove that if the diagonal sum of a product $y=x_1x_2\cdots x_m$ be finite, then it is unchanged by cyclic rearrangement

[1] As is known, a^{-1} is uniquely defined by (7) for *finite* square matrices when the determinant A of the matrix a is non-zero. If $A=0$ there is no matrix reciprocal to a.

of the factors:

$$D(x_1x_2 \cdots x_m) = D(x_rx_{r+1} \cdots x_mx_1x_2 \cdots x_{r-1}). \tag{10}$$

Clearly, it suffices to establish the validity of this rule for *two* factors.

If the elements of the matrices a and b are functions of a parameter t, then

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

or from the definition (3):

$$\frac{\mathrm{d}}{\mathrm{d}t}(ab) = \dot{a}b + a\dot{b}. \tag{11}$$

Repeated application of (11) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}(x_1x_2 \cdots x_n) = \dot{x}_1x_2 \cdots x_n + x_1\dot{x}_2 \cdots x_n + \ldots + x_1x_2 \cdots \dot{x}_n. \tag{11'}$$

From the definitions (2) and (3) we can define *functions* of matrices. To begin with, we consider as the most general function of this type, $f(x_1, x_2, \ldots x_m)$, one which can formally be represented as a sum of a finite or infinite number of products of powers of the arguments x_k weighted by numerical coefficients.

Through the equations

$$\begin{aligned} &f_1(y_1, \ldots y_n; x_1, \ldots x_n) = 0, \\ &\cdots\cdots\cdots\cdots\cdots \\ &f_n(y_1, \ldots y_n; x_1, \ldots x_n) = 0 \end{aligned} \tag{12}$$

we can then also define functions $y_l(x_1, \ldots x_n)$; namely, in order to obtain functions y_l having the above form and satisfying equation (12), the y_l need only be set in form of a series in increasing power products of the x_k and the coefficients determined through substitution in (12). It can be seen that one will always derive as many equations as there are unknowns. Naturally, the number of equations and unknowns exceeds that which would ensue from applying the method of undetermined coefficients in the normal type of analysis incorporating *commutative* multiplication. In each of the equations (12), upon substituting the series for the y_l and gathering together

like terms one obtains not only a sum term $C'x_1x_2$ but also a term $C''x_2x_1$ and thereby has to bring both C' and C'' to vanish (e.g., not only $C'+C''$). This is, however, made possible by the fact that in the expansion of each of the y_l, two terms x_1x_2 and x_2x_1 appear, with two available coefficients.

2. Symbolic differentiation

At this stage we have to examine in detail the process of *differentiation* of a matrix function, which will later be employed frequently in calculation. One should at the outset note that only in a few respects does this process display similarity to that of differentiation in ordinary analysis. For example, the rules for differentiation of a product or of a function of a function here no longer apply in general. Only if all the matrices which occur *commute* with one another can one apply all the rules of normal analysis to this differentiation.

Suppose

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

We define

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

This rule may be expressed as follows: In the given product, one regards all factors as written out *individually* (e.g., not as $x_1^3x_2^2$, but as $x_1x_1x_1x_2x_2$); one then picks out any factor x_k and builds the product of all the factors which follow this and which precede (in this sequence). The sum of all such expressions is the differential coefficient of the product with respect to this x_k.

The procedure may be illustrated by some examples:

$$y = x^n, \qquad \frac{dy}{dx} = nx^{n-1}$$

$$y = x_1^n x_2^m, \qquad \frac{\partial y}{\partial x_1} = x_1^{n-1}x_2^m + x_1^{n-2}x_2^m x_1 + \ldots + x_2^m x_1^{n-1},$$

$$y = x_1^2 x_2 x_1 x_3, \qquad \frac{\partial y}{\partial x_1} = x_1x_2x_1x_3 + x_2x_1x_3x_1 + x_3x_1^2x_2.$$

If we further stipulate that

$$\frac{\partial(y_1 + y_2)}{\partial x_k} = \frac{\partial y_1}{\partial x_k} + \frac{\partial y_2}{\partial x_k}, \tag{15}$$

then the derivative $\partial y/\partial x$ is defined for the most general analytical functions y.

With the above definitions, together with that of the diagonal sum (9), there follows the relation

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

on the right-hand side of which stands the mn-component of the matrix $\partial y/\partial x_k$. This relation can also be used to define the derivative $\partial y/\partial x_k$. In order to prove (16), it obviously suffices to consider a function y having the form (13). From (14) and (3) it follows that

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

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

On the other hand, from (3) and (9) ensues

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

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

Comparison of (17) with (17′) yields (16).

We here pick out a fact which will later assume importance and which can be deduced from the definition (14): *the partial derivatives of a product are invariant with respect to cyclic rearrangement of the factors*. Because of (16) this can also be inferred from (10).

To conclude this introductory section, some additional description is devoted to functions $g(pq)$ of *two* variables. For

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

it follows from (14) that

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

The most general function $g(pq)$ to be considered is to be represented in accordance with § 1 by a linear aggregate of terms

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

With the abbreviation

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

one can write the derivatives as

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

From these equations we find an important consequence. We consider the matrices

$$d_1 = q \frac{\partial z}{\partial q} - \frac{\partial z}{\partial q} q, \qquad d_2 = p \frac{\partial z}{\partial p} - \frac{\partial z}{\partial p} p. \tag{22}$$

From (21) we have

$$d_1 = \sum_{l=1}^{k} (q^{r_l} P_l p^{s_l} - P_l p^{s_l} q^{r_l}),$$

$$d_2 = \sum_{l=1}^{k} (p^{s_l} q^{r_l} P_l - q^{r_l} P_l p^{s_l}),$$

and thus it follows that

$$d_1 + d_2 = \sum_{l=1}^{k} (p^{s_l} q^{r_l} P_l - P_l p^{s_l} q^{r_l}).$$

Herein the second member of each term cancels the first member of the following, and the first and last member of the overall sum also cancel, so that

$$d_1 + d_2 = 0. \tag{23}$$

Because of its linear character in z, this relation holds not only for

expressions z having the form (19), but indeed for arbitrary analytical functions $g(pq)$.[1]

In concluding this brief survey of matrix analysis, we establish the following rule: *Every matrix equation*

$$F(x_1, x_2, \ldots x_r) = 0$$

remains valid if in all the matrices x_j one and the same permutation of all rows and columns is undertaken. To this end, it suffices to show that for two matrices a, b which thereby become transposed to a', b', the following invariance conditions apply:

$$a' + b' = (a + b)', \qquad a'b' = (ab)',$$

wherein the right-hand sides denote those matrices which are formed from $a+b$ and ab respectively by such an interchange.

We set forth this proof by replacing the procedure of permutation by that of multiplication with a suitable matrix.[2]
We write a permutation as

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

and to this we assign a *permutation matrix,*

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

The transposed matrix to $\mathfrak{p}$ is

$$\tilde{\mathfrak{p}} = (\tilde{p}(nm)), \qquad \tilde{p}(nm) = \begin{cases} 1 \text{ when } n = k_m \\ 0 \text{ otherwise.} \end{cases}$$

[1] More generally, for functions of r variables, one has

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

[2] The method of proof adopted here possesses the merit of revealing the close connection of permutations with an important class of more general transformations of matrices. The validity of the rule in question can however also be established directly on noting that in the definitions of *equality*, as also of *addition* and *multiplication* of matrices, no use was made of order relationships between the rows or the columns.

On multiplying the two together, one has

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

since the two factors $p(nk)$ and $\tilde{p}(km)$ differ from zero simultaneously only if $k=k_n=k_m$, i.e., when $n=m$. Hence $\tilde{p}$ is reciprocal to p:

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

If now a be any given matrix, then

$$pa = \left(\sum_k p(nk)\,a(km)\right) = \left(a(k_n, m)\right)$$

is a matrix which arises from the permutation $\binom{n}{k_n}$ of the rows of a, and equivalently

$$ap^{-1} = \left(\sum_k a(nk)\,\tilde{p}(km)\right) = \left(a(n, k_m)\right)$$

is the matrix arising from permutation of the columns of a. One and the same permutation applied both to the rows and the columns of a thus yields the matrix

$$a' = pap^{-1}.$$

Thence follows directly

$$a' + b' = p(a + b)p^{-1} = (a + b)',$$

$$a'b' = pabp^{-1} \qquad = (ab)',$$

which proves our original contention.

It is thus apparent that from matrix equations one can never determine any given sequence or order of rank of the matrix elements.

Moreover, it is evident that a much more general rule applies, namely that every matrix equation is invariant with respect to transformations of the type

$$a' = bab^{-1},$$

where b denotes an *arbitrary* matrix. We shall see later that this does not necessarily always apply to matrix differential equations.

CHAPTER 2. DYNAMICS

3. The basic laws

The dynamic system is to be described by the spatial coordinate q and the momentum p, these being represented by matrices

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

Here the $\nu(nm)$ denote the quantum-theoretical frequencies associated with transitions between states described by the *quantum numbers* n and m. The matrices (24) are to be Hermitian, e.g., on transposition of the matrices, each element is to go over into its complex conjugate value, a condition which should apply for all real t. We thus have

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

and

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

If q be a *Cartesian* coordinate, then the expression (25) is a measure of the *probabilities*[1] of the transitions $n \rightleftarrows m$.

Further, we shall require that

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

This can be expressed together with (26) in the following manner: there exist quantities W_n such that

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

From this, with equations (2), (3), it follows that a function $g(pq)$ invariably again takes on the form

$$g = (g(nm)e^{2\pi i\nu(nm)t}) \tag{29}$$

and the matrix $(g(nm))$ therein results from identically the same process applied to the matrices $(q(nm))$, $(p(nm))$ as was employed to find g from q, p. For this reason we can henceforth abandon the representation (24) in favour of the shorter notation

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

For the *time derivative* of the matrix $g = (g(nm))$, recalling to mind

[1] In this connection see § 8.

(24) or (29), we obtain the matrix

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

If $\nu(nm)\neq 0$ when $n\neq m$, a condition which we wish to assume, then the formula $\dot{g}=0$ denotes that g is a diagonal matrix with $g(nm)=\delta_{nm}g(nn)$.

A matrix differential equation $\dot{g}=a$ is invariant with respect to that process in which the same permutation is carried out on rows and columns of all the matrices and also upon the numbers W_n. In order to realize this, consider the diagonal matrix

$$W = (\delta_{nm}W_n).$$

Then

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

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

i.e., according to (31),

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

If now p be a permutation matrix, then the transform of W,

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

is the diagonal matrix with the permuted W_n along the diagonal. Thence one has

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

where $g'=pgp^{-1}$ and $\dot{g}'$ denotes the time derivative of g' constructed in accordance with the rule (31) with permuted W_n.

The rows and columns of $\dot{g}$ thus experience the same permutation as those of g, and hence our contention is vindicated.

It is to be noted that a corresponding rule does *not* apply to arbitrary transformations of the form $a'=bab^{-1}$ since for these W' is no longer a diagonal matrix. Despite this difficulty, a thorough study of these general transformations would seem to be called for, since it offers promise of insight into the deeper connections intrinsic to this new theory: we shall later revert to this point.[1]

[1] Cf. the continuation of this work, to be published forthwith.

In the case of a Hamilton function having the form

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

we shall assume, as did Heisenberg, that the *equations of motion* are just of the same form as in classical theory, so that using the notation of § 2 we can write:

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

We now use correspondence considerations to try more generally to elucidate the equations of motion belonging to an arbitrary Hamilton function $H(pq)$. This is required from the standpoint of relativistic mechanics and in particular for the treatment of electron motion under the influence of magnetic fields. For in this latter case, the function H cannot in a Cartesian coordinate system any longer be represented by the sum of two functions of which one depends only on the momenta and the other on the coordinates.

Classically, equations of motion can be derived from the action principle

$$\int_{t_0}^{t_1} L \, \mathrm{d}t = \int_{t_0}^{t_1} \{p\dot{q} - H(pq)\} \, \mathrm{d}t = \text{extremum}. \tag{33}$$

If we now envisage the Fourier expansion of L substituted in (33) and the time interval $t_1 - t_0$ taken sufficiently large, we find that only the constant term of L supplies a contribution to the integral. The form which the action principle thence acquires suggests the following translation into quantum mechanics:

The diagonal sum $D(L) = \sum_k L(kk)$ *is to be made an extremum:*

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

namely, by suitable choice of p *and* q, *with* $\nu(nm)$ *kept fixed.*

Thus, by setting the derivatives of $D(L)$ with respect to the elements

of p and q equal to zero, one obtains the equations of motion

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

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

From (26), (31) and (16) one observes that these equations of motion can always be written in *canonical* form,

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

For the quantization condition, Heisenberg employed a relation proposed by Thomas[1] and Kuhn.[2] The equation

$$J = \oint p\, \mathrm{d}q = \int_0^{1/\nu} p\dot{q}\, \mathrm{d}t$$

of 'classical' quantum theory can, on introducing the Fourier expansions 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 transformed into

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

If therein one has $p=m\dot{q}$, one can express the p_τ in terms of q_τ and thence obtain that classical equation which on transformation into a difference equation according to the principle of correspondence yields the formula of Thomas and Kuhn. Since here the assumption that $p=m\dot{q}$ should be avoided, we are obliged to translate equation (36) directly into a difference equation.

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

The following expressions should correspond:

$$\sum_{\tau=-\infty}^{\infty} \tau \frac{\partial}{\partial J}(q_\tau p_{-\tau}) \quad \text{with}$$

$$\frac{1}{h}\sum_{\tau=-\infty}^{\infty} q(n+\tau, n)p(n, n+\tau) - q(n, n-\tau)p(n-\tau, n)\big);$$

where in the right-hand expression those $q(nm)$, $p(nm)$ which take on a negative index are to be set equal to zero. In this way we obtain the quantization condition corresponding to (36) as

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

This is a system of infinitely many equations, namely one for each value of n.

In particular, for $p=m\dot{q}$ this yields

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

which, as may easily be verified, agrees with Heisenberg's form of the quantization condition, or with the Thomas–Kuhn equation. The formula (37) has to be regarded as the appropriate generalization of this equation.

Incidentally one sees from (37) that the diagonal sum $D(pq)$ necessarily becomes infinite. For otherwise one would have $D(pq)-D(qp)=0$ from (10), whereas (37) leads to $D(pq)-D(qp)=\infty$. Thus the matrices under consideration are never finite.[1]

4. Consequences. Energy-conservation and frequency laws

The content of the preceding paragraphs furnishes the basic rules of the new quantum mechanics in their entirety. All other laws of quantum mechanics, whose general validity is to be verified, must be *derivable* from these basic tenets. As instances of such laws to be proved, the law of energy conservation and the Bohr frequency condition primarily enter into consideration. The law of conservation of energy states that if H be the energy, then $\dot{H}=0$, or that H is a

[1] Further, they do not belong to the class of 'bounded' infinite matrices hitherto almost exclusively investigated by mathematicians.

diagonal matrix. The diagonal elements $H(nn)$ of H are interpreted, according to Heisenberg, as the *energies of the various states of the system* and the Bohr frequency condition requires that

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

or

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

We consider the quantity

$$d = pq - qp.$$

From (11), (35) one finds

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

Thus from (22), (23) it follows that $\dot{d}=0$ and d is a diagonal matrix. The diagonal elements of d are, however, specified just by the quantum condition (27). Summarizing, we obtain the equation

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

on introducing the unit matrix $\mathbf{1}$ defined by (6). We term the equation (38) the 'stronger quantum condition' and base all further conclusions upon it.

From the form of this equation, we deduce the following: If an equation (A) be derived from (38), then (A) remains valid if p be replaced by q and simultaneously h by $-h$. For this reason one need for instance derive only one of the following two equations from (38), which can readily be performed by induction

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

$$q^n p = pq^n - n\frac{h}{2\pi i}q^{n-1}. \tag{39'}$$

We shall now prove the energy-conservation and frequency laws, as expressed above, in the first instance for the case

$$H = H_1(p) + H_2(q).$$

From the statements of § 1, it follows that we may formally replace $H_1(p)$ and $H_2(q)$ by power expansions

$$H_1 = \sum_s a_s p^s, \qquad H_2 = \sum_s b_s q^s.$$

Formulae (39) and (39′) indicate that

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

Comparison with the equations of motion (35) yields

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

Denoting the matrix $Hg - gH$ by $|{}^{H}_{g}|$ for brevity, one has

$$\left|\begin{matrix} H \\ ab \end{matrix}\right| = \left|\begin{matrix} H \\ a \end{matrix}\right| b + a \left|\begin{matrix} H \\ b \end{matrix}\right|; \tag{42}$$

from which generally for $g = g(pq)$ one may conclude that

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

To establish this result, one need only conceive $\dot{g}$ as expressed in function of p, q and $\dot{p}$, $\dot{q}$ with the aid of (11), (11′), and $|{}^{H}_{g}|$ as evaluated by means of (42) in function of p, q and $|{}^{H}_{p}|$, $|{}^{H}_{q}|$, followed by application of the relations (41). In particular, if in (43) one sets $g = H$, one obtains

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

Now that we have verified the energy-conservation law and recognized the matrix H to be diagonal, equation (41) can be put into 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 follows.

If we now go over to consideration of more general Hamilton functions $H^*=H^*(pq)$, it can easily be seen that in general $\dot{H}^*$ no longer vanishes (examples such as $H^*=p^2q$, readily reveal this). It can however be observed that the Hamilton function $H=\frac{1}{2}(p^2q+qp^2)$ yields the same equations of motion as H^* and that $\dot{H}$ again vanishes. In consequence we may express the energy-conservation and frequency laws in the following way: *To each function* $H^*=H^*(pq)$ *there can be assigned a function* $H=H(pq)$ *such that as Hamiltonians* H^* *and* H *yield the same equations of motion and that for these equations of motion* H *assumes the role of an energy which is constant in time and which fulfils the frequency condition.*

On bearing in mind the considerations discussed above, it suffices to show that the function H to be specified satisfies not only the conditions

$$\frac{\partial H}{\partial p} = \frac{\partial H^*}{\partial p}, \qquad \frac{\partial H}{\partial q} = \frac{\partial H^*}{\partial q}, \tag{45}$$

but in addition satisfies equations (40). From § 1, the matrix H^* is formally to be represented as a sum of products of powers of p and q. Because of the linearity of equations (40), (45) in H, H^* we have simply to specify the commensurate sum term in H as counterpart to each individual sum term in H^*. Thus we need consider solely the case

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

It follows from the remarks of § 2 that equations (45) can be satisfied by specifying H as a linear form of those products of powers of p, q which arise from H^* through cyclic interchange of the factors; herein the sum of the coefficients must be held to unity. The question as to how these coefficients are to be chosen so that equations (40) may also be satisfied is less easy to answer. It may at this juncture suffice to dispose of the case $k=1$, namely

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

The formula (39) can be generalized[1] to

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

For $n=1$ this reverts to (39); in general (48) ensues from the fact that because of (39) one has

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

The new formula

$$p^m q^n - q^n p^m = n \frac{h}{2\pi i} \sum_{j=0}^{m-1} p^{m-1-j} q^{n-1} p^j \tag{48'}$$

is obtained on interchanging p and q and reversing the sign of h.

Comparison with (48) yields

$$\frac{1}{s+1} \sum_{l=0}^{s} p^{s-l} q^r p^l = \frac{1}{r+1} \sum_{j=0}^{r} q^{r-j} p^s q^j. \tag{49}$$

We now assert: The matrix H belonging to H^* as given by (47) is:

$$H = \frac{1}{s+1} \sum_{l=0}^{s} p^{s-l} q^r p^l. \tag{50}$$

We need only prove equations (40), to which end we recall the derivatives, (18') § 2.

From (50), we now obtain the relation

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

and according to (48) this is equivalent to the lower of equations (40).

Further, using (49) we find

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

[1] A different generalization is furnished by the formulae

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

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

where j runs to the lesser of the two integers m, n.

and by (48′) this is equivalent to the upper of equations (40). This completes the requisite proof.

Whereas in classical mechanics energy conservation ($\dot{H}=0$) is directly apparent from the canonical equations, the same law of energy conservation in quantum mechanics, $\dot{H}=0$ lies, as one can see, more deeply hidden beneath the surface.

That its demonstrability from assumed postulates is far from being trivial will be appreciated if, following more closely the classical method of proof, one sets out to prove H to be constant simply by evaluating $\dot{H}$. To this end, one first has to express $\dot{H}$ as function of p, q and $\dot{p}$, $\dot{q}$ with the aid of (11), (11′), whereupon for $\dot{p}$ and $\dot{q}$ the values $-\partial H/\partial q$, $\partial H/\partial p$ have to be introduced. This yields $\dot{H}$ in function of p and q. Equation (38) or the formulae quoted in the footnote to equation (48) which were derived from (38) permit this function to be converted into a sum of terms of the type ap^sq^r and one then has to prove that the coefficient a in each of such terms vanishes. This calculation for the most general case, as considered above along different lines, becomes so exceedingly involved[1] that it seems hardly feasible. The fact that nonetheless energy-conservation and frequency laws could be proved in so general a context would seem to us to furnish strong grounds to hope that this theory embraces truly deep-seated physical laws.

In conclusion, we append a result here which can easily be derived from the formulae of this section, namely: *Equations (35), (37) can be replaced by (38) and (44) (with H representing the energy); the frequencies are thereby to be derived from the frequency condition.*

In the continuation to this paper, we shall examine the important applications to which this theorem gives rise.

CHAPTER 3. INVESTIGATION OF THE ANHARMONIC OSCILLATOR

The anharmonic oscillator, having

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

has already been considered in detail by Heisenberg. Nevertheless, its

[1] For the case $H=(1/2m)p^2+U(q)$ it can immediately be carried out with the aid of (39′).

investigation will here be renewed with the aim of determining the *most general* solution of the fundamental equations for this case. If the basic equations of the present theory are indeed complete and do not require to be supplemented any further, then the absolute values $|q(nm)|$, $|p(nm)|$ of the elements of the matrices q and p must *uniquely* be determined by these equations, and thus it becomes important to check this for the example (51). On the other hand, it is to be expected that an uncertainty will still persist with respect to the phases φ_{nm}, ψ_{nm} in the relations

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

For the statistical theory, e.g., of the interaction of quantized atoms with external radiation fields, it becomes of fundamental importance to ascertain the precise degree of such uncertainty.

5. Harmonic oscillator

The starting point in our considerations is the theory of the harmonic oscillator; for small λ, one can regard the motion as expressed by equation (51) to be a perturbation of the normal harmonic oscillation having energy

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

Even for this simple problem it is necessary to supplement Heisenberg's analysis. This latter employs correspondence considerations to arrive at significant deductions as to the form of the solution: namely, since classically only a *single* harmonic component is present, Heisenberg selects a matrix which represents transitions between adjacent states only, and which thus has the form

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

We here strive to build up the entire theory self-dependently, without invoking assistance from classical theory on the basis of the principle of correspondence. We shall therefore investigate whether the form of the matrix (53) cannot itself be derived from the basic formulae or, if this proves impossible, which additional postulates are required.

From what has been stated in § 3 regarding the invariance with respect to permutation of rows and columns, one can see right away that the exact form of the matrix (53) can never be deduced from the fundamental equations, since if rows and columns be subjected to the same permutation, the canonical equations and the quantum condition remain invariant and thereby one obtains a new and apparently different solution. But all such solutions naturally differ only in the notation, i.e., in the way the elements are numbered. We seek to prove that through a mere renumbering of its elements, the solution can always be brought into the form (53). The equation of motion

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

runs as follows for the elements:

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

where

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

From the stronger quantum condition

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

it follows that for each n there must exist a corresponding n' such that $q(nn') \neq 0$, since if there were a value of n for which all $q(nn')$ were equal to zero, then the nth diagonal element of $pq - qp$ would be zero, which contradicts the quantum condition. Hence equation (55) implies that there is always an n' for which

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

But since we have assumed in our basic principles that when $n \neq m$, the energies are always unequal ($W_n \neq W_m$), it follows that at most *two* such indices n' and n'' can exist, for the corresponding $W_{n'}$, $W_{n''}$ are solutions of the quadratic equation

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

and if indeed *two* such indices n', n'' exist, it follows that the corresponding frequences must be related as:

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

Now from (56) we get

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

and the energy (52) ensues as

$$H(nm) = \tfrac{1}{2} \times 4\pi^2 \sum_k \{-\nu(nk)\,\nu(km)\,q(nk)\,q(km) + \nu_0^2 q(nk)\,q(km)\}$$

$$= 2\pi^2 \sum_k q(nk)\,q(km)\{\nu_0^2 - \nu(nk)\,\nu(km)\}.$$

In particular, for $m=n$ we have

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

Moreover, we can now distinguish between three possible cases:
(a) no n'' exists and one has $W_{n'} > W_n$;
(b) no n'' exists and one has $W_{n'} < W_n$;
(c) n'' exists.

In case (b) we now consider n' in place of n; to this there belong at most two indices $(n')'$ and $(n')''$ and of these, one has to equal n. We thereby revert to one of the cases (a) or (c) and can accordingly omit further consideration of (b).

In case (a), $\nu(n'n) = +\nu_0$ and from (58) it follows that

$$\nu_0|q(nn')|^2 = h/8\pi^2, \tag{60}$$

and thus from (59) that

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

Because of the assumption that $W_n \neq W_m$ for $n \neq m$ there is thus at most *one* index $n=n_0$ for which the case (a) applies.

If such an n_0 exists, we can specify a series of numbers n_0, n_1, n_2, n_3, ..., such that $(n_k)' = n_{k+1}$ and $W_{k+1} > W_k$. Then invariably $(n_{k+1})'' = n_k$. Hence for $k>0$, equations (58) and (59) give

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

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

From (60) and (62) it follows that

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

and thence from (61) that

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

Now, we still have to check whether it be possible that there is no value of n for which case (a) applies. Beginning with an arbitrary n_0 we can then build $n_0' = n$, and $n_0'' = n_{-1}$ and with each of these latter write $n_1' = n_2$, $n_1'' = n_0$ and $n_{-1}' = n_0$, $n_{-1}'' = n_{-2}$ etc. In this manner we obtain a series of numbers $\ldots n_{-2}, n_{-1}, n_0, n_1, n_2 \ldots$, and equations (61), (62) hold for every k between $-\infty$ and $+\infty$. But this is impossible, since by (62) the quantities $x_k = |q(n_{k+1}, n_k)|^2$ form an equispaced series of numbers, and since they are positive, there must be a least value. The relevant index can then again be designated as n_0 and we thereby revert to the previous case – thus here also, the formulae (63), (64) apply.

One can further see that every number n must be contained within the numbers n_k, since otherwise one could construct a new series (65) proceeding from n, and for this formula (60) would again hold. The starting terms of both series would then have the same value $W_n = H(nn)$, which is not possible.

This proves that the indices 0, 1, 2, 3 ... can be rearranged into a new sequence $n_0, n_1, n_2, n_3 \ldots$ such that formulae (63), (64) apply: with these new indices, the solution then takes on Heisenberg's form (53). Hence this appears as the 'normal form' of the general solution. By virtue of (64), it possesses the property that

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

If, inversely, one stipulate that $W_n = H(nn)$ should always increase with n, then it necessarily follows that $n_k = k$; this principle thus uniquely establishes the normal form of the solution. But thereby only the notation becomes fixed and the calculation more transparent: nothing new is conferred *physically*.

Therein lies the big difference between this and the previously adopted semiclassical methods of determining the stationary states. The classically calculated orbits merge into one another continuously; consequently the quantum orbits selected at a later stage have a particular sequence right from the outset. The new mechanics presents itself as an essentially discontinuous theory in that herein there is no question of a sequence of quantum states defined by the physical process, but rather of quantum numbers which are indeed no more

than distinguishing indices which can be ordered and normalized according to any practical standpoint whatsoever (e.g., according to increasing energy W_n).

6. Anharmonic oscillator

The equations of motion

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

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

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

We introduce series expansions

$$\begin{aligned} \omega(nm) &= \omega^0(nm) + \lambda\omega^{(1)}(nm) + \lambda^2\omega^{(2)}(nm) + \ldots \\ q(nm) &= q_0(nm) + \lambda q^{(1)}(nm) + \lambda^2 q^{(2)}(nm) + \ldots \end{aligned} \tag{68}$$

in seeking the solution.

When $\lambda=0$, one has the case of the harmonic oscillator considered in the previous section; we write the solution (53) in the form

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

where the bar denotes the conjugate complex value. If one builds the square or higher powers of the matrix $q^0=(q^0(nm))$, one arrives at matrices of similar form, being composed of sums of terms

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

This prompts us to try a solution of the form

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

n which odd and even values of the index p always alternate. If one

actually inserts this in the approximation equations

$$\lambda:\left\{\begin{aligned}&\left(\omega_0^2-\omega^0(nm)^2\right)q^{(1)}(nm)-2\omega^0(nm)\,\omega^{(1)}(nm)\,q^0(nm)\\&\qquad+\sum_k q^0(nk)\,q^0(km)=0,\\&\sum_k\{\omega^0(nk)\,\left(q^0(nk)\,q^{(1)}(kn)+q^{(1)}(nk)\,q^0(kn)\right)\\&\qquad+\omega^{(1)}(nk)\,q^0(nk)\,q^0(kn)\}=0,\end{aligned}\right\}\tag{72}$$

$$\lambda^2:\left\{\begin{aligned}&\left(\omega_0^2-\omega^0(nm)^2\right)q^{(2)}(nm)-2\omega^0(nm)\,\omega^{(1)}(nm)\,q^{(1)}(nm)\\&\qquad-\left(\omega^{(1)}(nm)^2+2\omega^0(nm)\,\omega^{(2)}(nm)\right)q^0(nm)\\&\qquad+\sum_k\left(q^0(nk)\,q^{(1)}(km)+q^{(1)}(nk)\,q^0(km)\right)=0,\\&\sum_k\{\omega^0(nk)\,\left(q^0(nk)\,q^{(2)}(km)+q^{(1)}(nk)\,q^{(1)}(km)\right.\\&\qquad\left.+q^{(2)}(nk)\,q^0(km)\right)+\omega^{(1)}(nk)\,\left(q^0(nk)\,q^{(1)}(km)\right.\\&\qquad\left.+q^{(1)}(nk)\,q^0(km)\right)+\omega^{(2)}(nk)\,q^0(nk)\,q^0(km)\}=0\end{aligned}\right\}\tag{73}$$

and notes the multiplication rule

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

one sees, in setting each of the factors of $\delta_{n,\,m-s}$ singly to zero, that through the substitution (71) all conditions can in fact be satisfied and that higher terms in (71) would identically vanish.

In detail, the calculation yields the following:

The first of the equations (72) gives, after substitution of the expressions (71),

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

and the second is identically satisfied. One thus has

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

The first of the equations (73) yields

$$\left.\begin{aligned} 2\omega_0 a_n \omega^{(2)}_{n,n+1} + 2a_n x_{n+1} + 2a_n x_n + \bar{a}_{n-1}x'_{n-1} + \bar{a}_{n+1}x'_n = 0, \\ -8\omega_0^2 y'_n + a_n x'_{n+1} + a_{n+2}x'_n = 0, \\ \omega^{(1)}_{n,n-2} = 0, \end{aligned}\right\} \quad (77)$$

whereas the second equation is not identically satisfied, but furnishes a relation from which y_n can be determined:

$$a_n\bar{y}_n + \bar{a}_n y_n - a_{n-1}\bar{y}_{n-1} - \bar{a}_{n-1}y_{n-1} + 2|x'_n|^2 - 2|x'_{n-2}|^2 - \frac{\omega^{(2)}_{n,n+1}}{\omega_0}|a_n|^2 - \frac{\omega^{(2)}_{n,n-1}}{\omega_0}|a_{n-1}|^2 = 0. \quad (78)$$

The solution is:

$$\left.\begin{aligned} \omega^{(2)}_{n,n+1} &= \frac{1}{3\omega_0^3}(|a_{n+1}|^2 + |a_{n-1}|^2 + 3|a_n|^2), \\ y'_n &= \frac{1}{12\omega_0^4} a_n a_{n+1} a_{n+2}. \end{aligned}\right\} \quad (79)$$

Further, if for brevity one introduces

$$\eta_n = a_n\bar{y}_n + \bar{a}_n y_n, \quad (80)$$

then the η are determined by the equation

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

Expressions (76) and (79) show that the quantities x_n, x'_n, y'_n can be expressed through the solution of the zero-th order approximation a_n. Thus their phases are determined by those of the harmonic oscillator. For the quantities y_n, the situation seems to be different, since although η_n can uniquely be determined from (81), y_n cannot be obtained absolutely from (80). It is probable that the next higher order of approximation gives rise to an auxiliary determining equation for y_n. We have to leave this question open here but we should like to indicate its significance as a point of principle in regard to the completeness of the entire theory. All questions of statistics invariably depend finally upon whether or not our supposition that of the phases of the $q(nm)$ *one* in each row (or each column) of the matrix remains undetermined be valid.

In conclusion we present the explicit formulae which are obtained by substituting the solution of the harmonic oscillator found previously (§ 5). In normal form, by (63), this runs as follows:

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

Thence, using (76), (79), (81) one obtains

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

$$\left.\begin{aligned} \omega^{(1)}_{n,n-1} &= 0, \qquad \omega^{(1)}_{n,n-2} = 0, \\ \omega^{(2)}_{n,n-1} &= -\frac{5C}{3\omega_2^3}n; \end{aligned}\right\} \tag{84}$$

that is,

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

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

If one sets $y_n = |y_n|e^{i\psi_n}$, then

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

In this approximation, y_n cannot be specified any more closely than this.

However, we should like to write out the final equations when one makes the assumption that $\psi_n = \varphi_n$. These are as follows (up to terms of higher than second order in λ):

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

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

We have also calculated the energy directly and derived the following formula:

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

The frequency condition is actually satisfied, since, remembering (82), we have

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

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

With the formula (88) we can associate the observation that already in terms of lowest order there occurs a discrepancy from classical theory which can formally be removed by the introduction of a 'half-integer' quantum number $n' = n + \frac{1}{2}$. This has already been remarked by Heisenberg. Incidentally, our expressions $\omega(n, n-1)$ as given by (86) agree *exactly* with the classical frequencies in all respects. For comparison, we note the classical energy to be[1]

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

and thus the classical frequency to be:

[1] See M. Born, Atommechanik (Berlin, 1925), Chapter 4, § 42, p. 294; one has to set $a = \frac{1}{3}$ in the formula (6) in order to obtain agreement with the present treatment.

$$\omega_{\mathrm{cl}} = \frac{1}{h}\frac{\partial W_n^{(\mathrm{cl})}}{\partial n} = h\nu_0 - \lambda^2 \frac{5C^2}{3\omega_0^2} n + \dots$$

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

We have, finally, checked that the expression (88) can also be derived from the Kramers–Born perturbation formula (up to an additive constant).