Received November 16, 1925

# 15

# ON QUANTUM MECHANICS II

**M. BORN, W. HEISENBERG AND P. JORDAN**

*Göttingen*

*Abstract:* The quantum mechanics developed in Part I of this paper from Heisenberg's approach is here extended to systems having arbitrarily many degrees of freedom. Perturbation theory is carried through for nondegenerate and for a large class of degenerate systems, and its connection with the eigenvalue theory of Hermitian forms is demonstrated. The results so obtained are employed in the derivation of momentum and angular momentum conservation laws, and of selection rules and intensity formulae. Finally, the theory is applied to the statistics of eigenvibrations of a black body cavity.

## Introduction

The present paper sets out to develop further a general quantum-theoretical mechanics whose physical and mathematical basis has been treated in two previous papers by the present authors.[1] It was found possible to extend the above theory to systems having several degrees of freedom[2] (Chapter 2), and by the introduction of 'canonical transformations' to reduce the problem of integrating the equations of motion to a known mathematical formulation. From this theory of canonical transformations we were able to derive a perturbation theory (Chapter 1, § 4) which displays close similarity to classical perturbation theory. On the other hand we were able to trace a connection between quantum mechanics and the highly-developed mathematical theory of quadratic forms of infinitely many variables (Chapter 3). Before we go on to

---

*Editor's note.* This paper was published as Z. Phys. **35** (1926) 557–615.

[1] W. Heisenberg, Zs. f. Phys. **33** (1925) 879.
M. Born and P. Jordan, Zs. f. Phys. **34** (1925) 858.
Henceforth designated as (Part) I.

[2] *Note added in proof:*
A paper by P. A. M. Dirac (Proc. Roy. Soc. London **109** (1925) 642), which has appeared in the meantime, independently gives some of the results contained in Part I and the present paper, together with further new conclusions to be drawn from the theory.

discuss the presentation of this further development in the theory, we first endeavour to define its physical content more precisely.

The starting point of our theoretical approach was the conviction that the difficulties which have been encountered at every step in quantum theory in the last few years could be surmounted only by establishing a mathematical system for the mechanics of atomic and electronic motions, which would have a unity and simplicity comparable with the system of classical mechanics, and which would entirely consist of relations between quantities that are in principle observable. Admittedly, such a system of quantum-theoretical relations between observable quantities, when compared with the quantum theory employed hitherto, would labour under the disadvantage of not being directly amenable to a geometrically visualizable interpretation, since the motion of electrons cannot be described in terms of the familiar concepts of space and time. A characteristic feature of the new theory lies in the modification it imposes upon kinematics as well as upon mechanics; a notable advantage, however, of this quantum mechanics consists in the fact that the basic postulates of quantum theory form an inherent organic constituent of this mechanics, e.g., that the existence of discrete stationary states is just as natural a feature of the new theory as, say, the existence of discrete vibration frequencies in classical theory (cf. Chapter 3). If one reviews the fundamental differences between classical and quantum theory, differences which stem from the basic quantum theoretical postulates, then the formalism proposed in the two above-mentioned publications and in this paper, if proved to be correct, would appear to represent a system of quantum mechanics as close to that of classical theory as could reasonably be hoped. In this context we merely recall the validity of energy and momentum conservation laws and the form of the equations of motion (Chapter 1, § 2). This similarity of the new theory with classical theory also precludes any question of a separate correspondence principle outside the new theory; rather, the latter can itself be regarded as an exact formulation of Bohr's correspondence considerations. In the further development of the theory, an important task will lie in the closer investigation of the nature of this correspondence and in the description of the manner in which symbolic quantum geometry goes over into visualizable classical geometry. With regard to this question, a particularly important trait in the new theory would seem to us to consist of the way in which both continuous and line spectra arise in it on an equal footing, i.e., as solutions of one and the same equation of motion and closely connected with one another mathematically (cf. Chapter 3, § 3); obviously, in this theory, any distinction between 'quantized' and 'unquantized' motion ceases to be at all meaningful, since the theory contains no mention of a quantization condition which selects only certain types of motion from among a large number of possible types: rather, in place of such a condition one has a basic quantum mechanical equation (Chapter 1, § 1) which is applicable to *all* possible types of motion and which is essential if the dynamic problem is to be given a definite meaning at all.

Now, although we should like to be able to conclude that because of its mathematical simplicity and unity, the proposed theory might reproduce essential characteristics of the actual conditions inherent in problems of atomic structure, we nevertheless have to realize, that the theory is not yet able to furnish a solution to the principal difficulties in quantum theory. The theory has not yet incorporated the forces which in classical theory would be associated with radiation resistance, and in connection with the question of how the coupling problem is to be related to the quantum mechanics postulated here, there exist but a few indistinct indications (cf. Chapter 1, § 5). Nevertheless it would seem that these basic quantum-theoretical difficulties assume an altogether different aspect in the new theory than hitherto and that one might indeed now be more justified in hoping that these problems will in due course be solved. We consider, for instance, the question of collision processes. Recently, Bohr[1] called attention to the basic difficulties which (in the theory as employed hitherto) confronted all attempts to reconcile the fundamental postulates of quantum theory with the law of conservation of energy in fast collisions. In the present theory, however, the fundamental principles of quantum theory and the principle of conservation of energy follow mathematically from the quantum-mechanical equations, and hence the results of the Franck–Hertz collision studies would seem to be natural mathematical consequences of the theory. One may thus hope that a future treatment of collision problems based on the new quantum mechanics may, just because of this organic

---

[1] N. Bohr, Zs. f. Phys. **34** (1925) 142.

relationship between the basic postulates and this mechanics, avoid difficulties of the type mentioned above.

The question of the anomalous Zeeman effect seems to be hardly different when handled by the theory proposed here than it was before. It is true that the intimate connection between the 'aperiodic' and the 'periodic' orbits inherent in the basic assumptions of this theory entails the fact that we cannot be certain that Larmor's Theorem holds generally (Chapter 4, § 2); the assumptions for the validity of the theorem are satisfied by an oscillator, but not necessarily by a nuclear atom. It is not likely, however, that this standpoint can lead to an interpretation of anomalous Zeeman effects; rather the present quantum mechanics may in the case of Zeeman effects have to content with the same difficulties as the previous theory. Recently, though, the problem of anomalous Zeeman effects has entered a new phase as a result of a Note published by Uhlenbeck and Goudsmit.[1] These authors make the assumption that the electron itself posseses a mechanical and a magnetic moment (whose ratio should be twice as large as for atoms), so that there should actually be no anomalous Zeeman effects. By this assumption, difficulties as to statistical weights are eliminated and a qualitative explanation of various phenomena connected with problems of multiplet structure and Zeeman effects ensues. The question as to whether it can already furnish a quantitative explanation of these phenomena can, of course, be answered only after more rigorous investigations using the methods of quantum mechanics. Some of the results contained in Chapter 4 appear, as regards the Zeeman effects, to substantiate this hope of finding a quantitative interpretation at some later date.

Finally, we have also attempted to treat a well-known statistical problem by means of the methods furnished by the present theory. It is well known that by quantizing the vibrations of a cavity within reflecting walls and using classical methods one can arrive at results which display a certain similarity with the hypotheses in a theory of light quanta and which permit a derivation of Planck's formula. However, as Einstein[2] has always stressed, this semiclassical treatment of cavity radiation yields an erroneous value for the mean square deviation of the energy in a volume element. This result must be

---

[1] G. Uhlenbeck and S. Goudsmit, Naturwiss. **13** (1925) 953.

[2] A. E instein, Phys. Zs. **10** (1909) 185, 817.

regarded as a particularly serious objection to earlier methods in quantum theory, since we are concerned here with a breakdown of the theory even for the simple problem of a harmonic oscillator. On the other hand, the above difficulty would arise in the statistical treatment of the eigenvibrations of any mechanical system whatsoever, e.g., a crystal lattice. Now, we have found that with the kinematics and mechanics inherent in the theory presented here, the corresponding calculation leads to a correct value for the mean square deviation and also to Planck's formula, a result which may well be regarded as significant evidence in favour of the quantum mechanics put forward here.

## CHAPTER 1. SYSTEMS HAVING ONE DEGREE OF FREEDOM

### 1. Fundamental principles

I. A quantum-theoretical quantity $a$, whether representing a coordinate or a momentum or any function of both, is depicted by a set of quantities

$$a(nm)e^{2\pi i\nu(nm)t} \tag{1}$$

or (on leaving off the factor $e^{2\pi i\nu(nm)t}$, which is the same for all quantities belonging to a given system and which depends only upon the indices $n$ and $m$) by the set of numbers

$$a(nm). \tag{2}$$

We can thus speak of an infinite 'matrix' $a$.

II. Elementary operations such as addition and multiplication of quantum-theoretical quantities are defined in accordance with the operational rules of matrix calculus.

III. Consider a given function $f(x_1, x_2, \ldots, x_s)$ defined through addition and multiplication of given matrices, with $x_1, x_2, \ldots, x_s$ denoting quantum-theoretical quantities. We then introduce two types of derivatives of $f$ with respect to one of the quantities $x$ (say, $x_1$):

(a) Differential coefficient of the first type:

$$\frac{\partial f}{\partial x_1} = \lim_{\alpha\to 0} \frac{f(x_1 + \alpha\mathbf{1}, x_2, \ldots, x_s) - f(x_1, x_2, \ldots, x_s)}{\alpha}, \tag{3}$$

where $\alpha$ represents a number and $\mathbf{1}$ the unit matrix defined by

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

(b) Differential coefficient of the second type: Defined through[1]

$$\frac{\partial f}{\partial x_1}(nm) = \frac{\partial D(f)}{\partial x_1(mn)}, \tag{4}$$

where $D(f)$ represents the diagonal sum of the matrix $f$.

These two forms of differentiation will be distinguished typographically by different fraction strokes [thick stroke for (a), thin for (b)].

The treatment in Part I employed differentiation of the *second* type exclusively since this leads to a simple formulation of the variational principle of quantum mechanics and hence appears to be the more natural. However, for some calculations derivatives of the first type are more convenient to employ. It might be mentioned generally that the introduction of a differential coefficient into quantum mechanics is somewhat of an artifice and that the operations on the left-hand side of the formula (6) which follow represent the natural counterpart to differential coefficients in classical theory. For the formulation of canonical equations it is important to establish the fact that both species of differentiation (3) and (4) become identical in the case of the energy function[2] $H(pq)$.

---

[1] Cf. Part I [paper **13** in this volume].

[2] For the energy function $H$ of Part I, instead of arbitrary functions such as

$$H^* = \sum a_{sr} p^s q^r,$$

only those symmetrized functions giving rise to the same Hamilton equations were permitted:

$$H = \sum a_{sr} \frac{1}{s+1} \sum_{l=0}^{s} p^{s-l} q^r p^l.$$

Now, for these symmetrized functions $H$ the following relations, derived in Part I, apply:

$$\begin{aligned} \frac{\partial H}{\partial p} &= \sum a_{sr} \frac{1}{s+1} \left\{ \sum_{l=0}^{s-1} (s-l) p^{s-1-l} q^r p^l + \sum_{l=1}^{s} l p^{s-l} q^r p^{l-1} \right\} \\ &= \sum a_{sr} \sum_{l=0}^{s-1} p^{s-1-l} q^r p^l = \frac{\partial H}{\partial p}. \\ \frac{\partial H}{\partial q} &= \sum a_{sr} \frac{r}{s+1} \sum_{l=0}^{s} p^{s-l} q^{r-1} p^l = \sum a_{sr} \sum_{j=0}^{r-1} q^{r-1-j} p^s q^j = \frac{\partial H}{\partial q}. \end{aligned}$$

IV. Calculations involving quantum-theoretical quantities would yield non-unique results because of the inapplicability of the commutative rule in multiplication unless the value of $pq-qp$ were prescribed.[1] Hence we introduce the following basic quantum-mechanical relation:

$$pq - qp = \frac{h}{2\pi \mathrm{i}} \mathbf{1}. \tag{5}$$

We shall later discuss the physical significance of this relation according to the correspondence principle. At this stage it would appear important to stress that eq. (5), ch. 1, is the only one of the basic formulae in the quantum mechanics here proposed which contains Planck's constant $h$. It is satisfying that the constant $h$ already enters into the basic tenets of the theory at this stage in so simple a form. Furthermore, one can see from eq. (5), ch. 1, that in the limit $h=0$, the new theory would converge to classical theory, as is physically required.

A relation which will later prove important can also be derived from eq. (5), ch. 1, namely:
If $f(pq)$ be any function of $p$ and $q$, then

$$\begin{aligned} fq - qf &= \frac{\partial f}{\partial p} \frac{h}{2\pi \mathrm{i}}, \\ pf - fp &= \frac{\partial f}{\partial q} \frac{h}{2\pi \mathrm{i}}, \end{aligned} \tag{6}$$

since, if we assume these formulae to be valid for some given pair of functions, $\varphi$ and $\psi$, then they must also hold for $\varphi+\psi$ and $\varphi\cdot\psi$. The former case, $\varphi+\psi$ is trivial; for the latter, $\varphi\cdot\psi$, a simple calculation yields:

$$\begin{aligned} \varphi\cdot\psi q - q\varphi\psi &= \varphi(\psi q - q\psi) + (\varphi q - q\varphi)\psi \\ &= \varphi\left(\frac{\partial\psi}{\partial p} + \frac{\partial\varphi}{\partial p}\psi\right)\frac{h}{2\pi \mathrm{i}} = \frac{\partial(\varphi\psi)}{\partial p}\ \frac{h}{2\pi \mathrm{i}}; \end{aligned}$$

for $p\varphi\psi-\varphi\psi p$, the treatment is similar.

Now, the relations (6) hold for $p$ and $q$. They must accordingly also apply to every function $f$ which can formally be expressed as a power series in $p$ and $q$.

---

[1] The equations of motion merely indicate that this difference has to be a diagonal matrix.

## 2. The canonical equations, energy conservation and frequency condition

Let an energy function $H(pq)$ be given, together with the associated canonical equations

$$\dot{p} = -\frac{\partial H}{\partial q}; \qquad \dot{q} = \frac{\partial H}{\partial p}. \tag{7}$$

It follows from the frequency combination principle

$$\nu(nm) + \nu(mk) = \nu(nk) \tag{8}$$

that $\nu$ can be expressed in the form

$$\nu(nm) = \frac{(W_n - W_m)}{h}. \tag{9}$$

We now introduce a quantum-theoretical quantity $W$, as 'term', defined through

$$W(nm) = \begin{cases} W_n & \text{for } n = m \\ 0 & \text{for } n \neq m. \end{cases}$$

Thus $W$ is a diagonal matrix.

Then for any quantum-theoretical quantity whatsoever, the following relation holds:

$$\dot{a} = \frac{2\pi \mathrm{i}}{h}(Wa - aW). \tag{10}$$

In fact $\dot{a}$ was (cf. Part I) defined through

$$a(\dot{n}m) = 2n\mathrm{i}\nu(nm)\, a(nm).$$

Among the main tenets of the theory we here seek to build up, we class the law of conservation of energy ($H$=constant) and the frequency condition

$$\left(\nu(nm) = \frac{H_n - H_m}{h}; \qquad H_n = W_n + \text{const}\right).$$

We carry the proof through for both these conditions by inserting eqs. (6) and (10) into eq. (7), ch. 1. This yields

$$\begin{aligned} Wq - qW &= Hq - qH \\ Wp - qW &= Hp - qH \end{aligned} \tag{11}$$

or, equivalently,

$$(W - H)q - q(W - H) = 0,$$
$$(W - H)p - q(W - H) = 0.$$

The entity $W{-}H$ commutes with $p$ and $q$, and hence also with every function of $p$, $q$, in particular with $H$:

$$(W - H)H - H(W - H) = 0.$$

Thence from (10), ch. 1, one has

$$\dot{H} = 0. \tag{12}$$

Thereby the law of conservation of energy is proved, and $H$ is established as a diagonal matrix, $H(nm) = \delta_{nm} H_n$.

The frequency condition now follows directly from (11), ch. 1:

$$q(nm)(H_n - H_m) = q(nm)(W_n - W_m), \tag{13}$$

i.e.,

$$\frac{(H_n - H_m)}{h} = \nu(nm). \tag{14}$$

Thus far, we have proved energy-conservation and the frequency condition from the canonical equations and the basic equation (5), ch. 1. In corollary, we can, however, also invert the proof. We know energy conservation and the frequency condition to be correct. Hence if the energy function $H$ be given as an analytical function of any variables $P$, $Q$ then, provided that

$$PQ - QP = \frac{h}{2\pi \mathrm{i}} \mathbf{1},$$

the following canonical equations always apply:

$$\dot{Q} = \frac{\partial H}{\partial P}, \qquad P = -\frac{\partial H}{\partial Q}. \tag{15}$$

This follows directly from the fact that the quantities $PH{-}HP$ or $HQ{-}QH$ can be interpreted in a twofold manner, namely according to (6), ch. 1 and according to (10), ch. 1.

## 3. Canonical transformations

By a 'canonical transformation' of the variables $p$, $q$ into new variables

$P$, $Q$, we understand a transformation in which

$$pq - qp = PQ - QP = \frac{h}{2\pi i}, \tag{16}$$

as is suggested by the preceding considerations, since then the same canonical equations (7), ch. 1, or (15), ch. 1, apply to $P$, $Q$ as to $p$ $q$.

A general transformation which satisfies this condition is

$$\begin{aligned} P &= SpS^{-1} \\ Q &= SqS^{-1}, \end{aligned} \tag{17}$$

wherein $S$ stands for an arbitrary quantum-theoretical quantity. We would surmise that eq. (17), ch. 1, represents in fact the *most general* canonical transformation. The transformation (17), ch. 1, also has the simple property that for any function $f(P, Q)$ it follows that

$$f(P, Q) = Sf(p, q)S^{-1}, \tag{18}$$

wherein $f(p, q)$ is formed from $f(P, Q)$ on replacing $P$ by $p$ and $Q$ by $q$, retaining the functional form. The proof of this contention for functions in the sense of our above definition follows directly from the observation that the rule holds for sum and product with sum terms or factors $p$, $q$.

The importance of the canonical transformation is due to the following theorem: If any pair of values $p_0$, $q_0$ be given which satisfy eq. (15), ch. 1, then the problem of integrating the canonical equations for an energy function $H(pq)$ can be reduced to the following: A function $S$ is to be determined, such that when

$$p = Sp_0S^{-1}, \qquad q = Sq_0S^{-1} \tag{19}$$

the function

$$H(pq) = SH(p_0q_0)S^{-1} = W \tag{20}$$

becomes a diagonal matrix. Equation (20), ch. 1, is the analogue to the Hamilton partial differential equation, and in a sense stands for the action function.

## 4. Perturbation theory

We consider a given mechanical problem defined by the energy function

$$H = H_0(pq) + \lambda H_1(pq) + \lambda^2 H_2(pq) + \dots \tag{21}$$

and assume the mechanical problem defined by the energy function $H_0(pq)$ to be solved. Thus solutions $p_0$, $q_0$ of this problem are known; they satisfy the condition $p_0q_0-q_0p_0=(h/2\pi i)\mathbf{1}$ and cause $H_0(p_0q_0)=W_0$ to be a diagonal matrix. We then seek a transformation function $S$ such that

$$p = Sp_0S^{-1}, \qquad q = Sq_0S^{-1}, \tag{22}$$

and that

$$H(pq) = SH(p_0q_0)S^{-1} = W,$$

e.g., that the matrix $H$ becomes diagonalized. To arrive at a solution we try setting

$$S = \mathbf{1} + \lambda S_1 + \lambda S_2 + \dots . \tag{23}$$

Then

$$S^{-1} = \mathbf{1} - \lambda S_1 + \lambda^2(S_1^2 - S_2) + \lambda^3 \dots . \tag{24}$$

If for $H$ we take the expression (21), ch. 1, we can collect together powers of $\lambda$ to obtain the following equations of approximation:

$$\begin{aligned} H_0(p_0q_0) &= W_0 \\ S_1H_0 - H_0S_1 + H_1 &= W_1 \\ S_2H_0 - H_0S_2 + H_0S_1^2 - S_1H_0S_1 + S_1H_1 - H_1S_1 + H_2 &= W_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ S_rH_0 - H_0S_r + F_r(H_0, \dots, H_r, S_0, \dots, S_{r-1}) &= W_r \end{aligned} \tag{25}$$

where $H_0$, $H_1$, ... are throughout to be taken as having arguments $p_0$, $q_0$.

The first of the eqs. (25), ch. 1, is already satisfied. The others can be resolved in sequence, actually in just the same manner as in classical theory, namely by first building the mean value in order to determine the energy constant, after which the solution can straightway be written down:

$$W_r = \bar{F}_r, \tag{26}$$

$$S_r(mn) = \frac{F_r(mn)}{h\nu_0(mn)}\,(1 - \delta_{nm}),$$

where $\nu_0(nm)$ are the frequencies of the unperturbed motion. This solution satisfies the condition

$$S \cdot \tilde{S}^* = \mathbf{1}, \tag{27}$$

wherein the tilde represents interchange of rows and columns (transposition) and the star denotes that we take the complex conjugate quantity. Since we shall later return to this condition from a more general standpoint we confine ourselves at this stage merely to verifying it to the first order of approximation, which we shall evaluate right away. To this order, the relation runs

$$S_1 + \tilde{S}_1^* = 0. \tag{28}$$

The significance of eq. (27), ch. 1, lies in the fact that the Hermitian character of the matrices $p$, $q$ follows from it, since use of (22), ch. 1, shows[1] that

$$q^* = S^* q_0^* S^{*-1} = \overset{*}{\tilde{S}}{}^{-1} \tilde{q}_0 \overset{*}{\tilde{S}} = \tilde{q},$$

and analogously for $p$.

To first approximation it follows from (26), ch. 1, as also classically, that

$$W_1 = \overline{H}_1, \tag{29}$$

so that

$$S_1(mn) = \frac{H_1(mn)}{h\nu_0(mn)} (1 - \delta_{mn}). \tag{30}$$

This expression indeed satisfies the requirements (28), ch. 1, because $H_1$ is assumed to be a Hermitian form. We can now evaluate the energy to the second order of approximation and find

$$W_2 = \overline{H}_2 + \frac{1}{h} \sum_l{}' \frac{H_1(nl) H_1(ln)}{\nu_0(nl)}, \tag{31}$$

where the prime on the summation indicates that terms having a vanishing denominator ($l = n$) are to be excluded.

One can progress in this way and successively determine all terms of the $W$ and $S$ series. If we substitute the $S$ series in (22), ch. 1, we obtain the expansions

$$q = q_0 + \lambda q_1 + \lambda^2 q_2 + \ldots,$$
$$p = p_0 + \lambda p_1 + \lambda^2 p_1 + \ldots$$

with known coefficients. Thus, for example, the first-order approximation runs

$$q_1 = S_1 q_0 - q_0 S_1,$$
$$p_1 = S_1 p_0 - p_0 S_1;$$

---

[1] **On noting the rule $\widetilde{(ab)} = \tilde{b}\tilde{a}$.**

or, explicitly,

$$\begin{aligned} q_1(mn) &= \frac{1}{h} \sum_k{}' \left( \frac{H_1(mk)q_0(kn)}{\nu_0(mk)} - \frac{q_0(mk)H_1(kn)}{\nu_0(kn)} \right) \\ p_1(mn) &= \frac{1}{h} \sum_k{}' \left( \frac{H_1(mk)q_0(kn)}{\nu_0(mk)} - \frac{q_0(mk)H_1(kn)}{\nu_0(kn)} \right). \end{aligned} \tag{32}$$

The formulae (32), ch. 1, represent the outcome of Kramers' dispersion theory[1] in the limit of an infinitely low-frequency external field; this possibility of attaining a simple derivation of formulae otherwise obtained only on the basis of correspondence considerations seems to provide a strong argument in favour of the theory put forward here. Born[2] has derived eq. (31), ch. 1, on reinterpreting the respective classical formulae. The terms with $m=n$ in eq. (32), ch. 1, correspond to Kramers' formula for normal dispersed light and the remaining terms $(m \neq n)$ correspond to the formulae of Kramers and Heisenberg[3] for 'scattered light of combination frequencies'. The latter expressions were used by Pauli[4] to evaluate the intensities of transitions in Hg which take place in presence of external electric fields and which would otherwise be 'forbidden'. In order to derive the general dispersion formulae (if the frequency of the external field does not vanish), one needs more general considerations regarding the action of external fields which change in function of time. We now pass over to such considerations.

## 5. Systems for which time-variables enter explicitly into the 'energy function'

Treatment of the quantum-mechanical influence of external forces which explicitly depend upon time seems to us to be of especial interest in that therein some characteristic differences crop up between classical and quantum mechanics. The problem of the action of time-dependent external forces can be regarded as a limiting case of the interaction between two systems in which the influence of the inter-

---

[1] H. A. Kramers, Nature **113** (1924) 673; **114** (1924) 310; cf. also R. Ladenburg, Zs. f. Phys. **4** (1921) 451; R. Ladenburg and F. Reiche, Naturwiss. **11** (1923) 584.

[2] M. Born, Zs. f. Phys. **26** (1924) 379.

[3] H. A. Kramers and W. Heisenberg, Zs. f. Phys. **31** (1925) 681.

[4] W. Pauli, Verh. d. Dän. Akad. d. Wiss. (in press).

action on one of the two systems (termed system $A$) is so small that the action upon the other system (system $B$) remains unaffected by this influence. If we now consider the coupling of two systems $A$, $B$ from the standpoint of quantum mechanics, the Hamilton function decomposes into three parts, $H_A$, $\lambda H_B$ and $\varepsilon\lambda H_{AB}$ (with $\lambda$ at this stage an arbitrary parameter and $\varepsilon$ a small quantity). We take system $A$ to be known. For calculating the motion of $B$ according to classical theory it suffices to establish the equations of motion [from the Hamilton function $\lambda(H_B+\varepsilon H_{AB})$] for the coordinates of $B$, whereby for the coordinates of $A$ one substitutes their solutions in function of time (for the definite given values of the constants in $A$). By this means, apart from the constants of $A$ only the time enters as a new variable into the perturbation problem for $B$ when the reaction is neglected. In the quantum-mechanical calculation the situation is just the same, providing we restrict ourselves to first-order perturbations (i.e., terms proportional to $\varepsilon$ in the coordinates and momenta of the system $B$). It is altogether otherwise, however, for higher-order perturbations, since in the evaluation of higher-order perturbations we encounter products of quantities in which more than one implicitly contains the coordinates of $A$. But this means that according to the quantum-mechanical rule for building a product it by no means suffices to know the 'external forces in function of time' merely for the *given* values of the constants in $A$, but these external forces must be known for *all* values of the constants. Thereby, however, the concept of external forces appears in fact to become devoid of meaning. This difficulty seems to us to be overcome on observing that the reaction itself gives rise to terms of order $\lambda\varepsilon^2$ in the coordinates of $B$, and thus that simultaneous neglect of the reaction and evaluation of terms in $B$ containing $\varepsilon^2$ is meaningful only if $\lambda$ can also be taken to be very small, i.e., physically, if variation of the quantities in $A$ by amounts of the same order as the associated quantities in $B$ does not bring about any perceptible change in the influence of $A$ upon $B$. However, in this approximation the quantum-mechanical construction of products and thereby the calculation of the perturbations to higher orders in $\varepsilon$ can again be effected. In fact, the rules for this building of products reduce simply to those of classical multiplication, as in this approximation the coordinates, amplitudes and frequencies which enter into $H_{AB}$ do not depend on the constants in $A$. In this sense one could, for example, treat the action of a strong alternating electromagnetic field on an atom entirely as the influence of an 'external force' with neglect of the reaction, since the field energy can be regarded as infinitely large compared with that of the atom. The action of $\alpha$-particles upon the electrons of an atom could also be regarded as an 'external force', as in classical theory, because of the relatively large energy of the $\alpha$-particles, so that in this approximation the Fourier expansion of the force thereby exerted upon the electrons would also be that of classical theory. However, the action of forces due to one atom upon another can never be treated as an operation of external forces – i.e., it can thus be regarded only in the first-order terms, for which such an approach is always possible – since the neglect of the reaction would in the higher-order terms lead to false results.

We can summarize the outcome of our considerations thus: It is meaningful under certain assumptions in quantum as in classical theory to speak of the action of time-dependent forces upon an atom. In such instances, the classical calculation rules can be applied to expressions in which the time parameter figures explicitly: e.g., if the external field of force be periodic with a period $\nu_0$, then the general term of a coordinate $q$ can be written as

$$q(mn, \tau)\, \mathrm{e}^{2\pi \mathrm{i}[\nu(mn)+\tau\nu_0]t}, \tag{33}$$

and the general term of $q^2$ as

$$\sum_{k,\tau'} q(mk, \tau - \tau')q(kn, \tau')\mathrm{e}^{2\pi \mathrm{i}[\nu(mn)+\tau\nu_0]t}. \tag{34}$$

For this reason the case of external forces which vary with time seems in our view to provide a striking illustration of the transition from theoretical quantum kinematics into classical kinematics according to the principle of correspondence.

If one is concerned with the evaluation of the operation of external forces to first order only, the results which ensue from the calculations which follow remain correct even if the assumptions listed at the outset are not obeyed – in exact analogy with the situation in classical theory.

From the preceding considerations it follows that the mathematical treatment of systems in which (provided the assumptions mentioned above are valid) time enters explicitly is simply to be handled in a manner analogous to the corresponding classical procedures. If we again assume the external force to be periodic in time, with period

$\nu_0$, the Hamilton function becomes[1]

$$H = H(p_k, q_k, \cos 2\pi\nu_0 t). \tag{35}$$

We then introduce a new degree of freedom with the variables $q'$, $p'$ and take the following as the Hamiltonian of the new problem, in which time no longer figures explicitly:

$$H' = H(p_k, q_k;\ q') + 2\pi\nu_0\sqrt{1 - q'^2}\,p'. \tag{36}$$

Thereby the canonical equations for $p_k$, $q_k$ remain as hitherto, except that $q'$ is throughout written for $\cos 2\pi\nu_0 t$. The new equations are:

$$\begin{aligned}\dot{q}' &= \frac{\partial H'}{\partial p'} = 2\pi\nu_0\sqrt{1 - q'^2},\\ \dot{p}' &= -\frac{\partial H'}{\partial q'} = -\frac{\partial H}{\partial q'} + 2\pi\nu_0\frac{q'}{\sqrt{(1 - q'^2)}}p'.\end{aligned} \tag{37}$$

The first of these equations asserts that $q'$ indeed becomes equal to $\cos 2\pi\nu_0 t$ (up to an arbitrary choice of origin in the time scale), so that the canonical equations for $p_k$, $q_k$ take on the same form as in the earlier problem; the second equation (37), ch. 1, provides a determination of $p'$. Thus through (36), ch. 1, the problem (35), ch. 1, is really led back to cases already treated.

Of paramount interest is the question as to the manner in which the perturbation formulae (25), ch. 1, have to be modified if time enters explicitly into $H_1, H_2, \ldots$ but *not* into $H_0$. Simple considerations show that for this case the perturbation formulae ensue from those cited earlier on replacing every term of the form $H_0S_r - S_rH_0$ by

$$H_0S_r - S_rH_0 + \frac{h}{2\pi i}\frac{\partial S_r}{\partial t}$$

(note that $H_0$ occurs only in such combinations). Thus the lowest orders of the new perturbation formulae run:

$$\begin{aligned}H_0(p_0q_0) &= W_0,\\ S_1H_0 - H_0S_1 - \frac{h}{2\pi i}\frac{\partial S_1}{\partial t} + H_1 &= W_1,\end{aligned} \tag{38}$$

---

[1] Here we anticipate for a moment in availing ourselves of results derived in the next chapter for systems having several degrees of freedom.

$$S_2H_0 - H_0S_2 - \frac{h}{2\pi i}\frac{\partial S_2}{\partial t} + \left(H_0S_1 - S_1H_0 + \frac{h}{2\pi i}\frac{\partial S_1}{\partial t}\right)S_1 \qquad (38)$$

$$+ S_1H_1 - H_1S_1 + H_2 = W_2,$$

. . . . . . . . . . . . . . . . . . . . . . . . . . .

We should like to assume that even if the assumption that the external forces are periodic in time does not apply, these formulae (38), ch. 1, nevertheless remain valid – even though this assumption was incorporated into the derivation of the formulae.

The first-order equations in the formulae (38), ch. 1, which of course remain correct even if the assumptions regarding 'external forces' are no longer valid, taken together with eqs. (22), ch. 1, viz.

$$q = q_0 + \lambda(S_1q_0 - q_0S_1),$$
$$p = p_0 + \lambda(S_1p_0 - p_0S_1),$$

furnish an answer to the problems of dispersion theory in a general sense. In actual fact, if we set:

$$H_1 = Eeq_0 \cos 2\pi\nu_0 t,$$

then

$$H_1(mn, 1) = \frac{Ee}{2}\, q_0(mn), \qquad H_1(mn, -1) = \frac{Ee}{2}\, q_0(mn),$$

$$S_1(mn, 1) = \frac{Ee}{2h}\, \frac{q_0(mn)}{\nu_0(mn) + \nu_0}, \qquad (39)$$

$$S_1(mn, -1) = \frac{Ee}{2h}\, \frac{q_0(mn)}{\nu_0(mn) - \nu_0}.$$

Thence follows (cf. (22), ch. 1):

$$q_1(mn, +1) = \frac{Ee}{2h} \sum_k \left(\frac{q_0(mk)q_0(kn)}{\nu_0(mk) + \nu_0} - \frac{q_0(mk)q_0(kn)}{\nu_0(kn) + \nu_0}\right). \qquad (40)$$

If we assume that we have Cartesian coordinates, i.e., $p = m\dot{q}$, then

$$q_1(mn, 1) = \frac{Ee}{2h\cdot 2\pi i m} \sum_k \frac{q_0(mk)p_0(kn) - p_0(mk)q_0(kn)}{(\nu_0(mk) + \nu_0)(\nu_0(kn) + \nu_0)}; \qquad (41)$$

and similarly

$$q_1(mn, -1) = \frac{Ee}{2h\cdot 2\pi i m} \sum_k \frac{q_0(mk)p_0(kn) - p_0(mk)q_0(kn)}{(\nu_0(mk) - \nu_0)(\nu_0(kn) - \nu_0)}. \qquad (42)$$

The eqs. (40), (41), (42), ch. 1 agree with the formulae obtained from Kramers' dispersion theory.[1] A further particularly interesting case would seem to be that for incident light of very high frequency, $|\nu_0| \gg |\nu_0(mk)|$ or $|\nu_0(kn)|$. Then to first-order approximation one finds

$$q_1 = -\frac{Ee}{h2\pi i\nu_0^2 m}(p_0q_0 - q_0p_0)\cos 2\pi\nu_0 t,$$

or, because of (5), ch. 1,

$$q_1 = +\frac{Ee}{4\pi^2 m\nu_0^2}\cos 2\pi\nu_0 t. \tag{43}$$

This finding indicates that in fact the quantum-mechanical commutation relation (5), ch. 1, ultimately entails the fact that for sufficiently high frequencies the electron behaves on scattering like a free electron. The scattered light of frequency $\nu_0(mn)+\nu_0(m\neq n)$ vanishes, and that of frequency $\nu_0$ has the intensity to be expected for scattering by a free electron.[2]

## CHAPTER 2. FUNDAMENTALS OF THE THEORY FOR SYSTEMS HAVING AN ARBITRARY NUMBER OF DEGREES OF FREEDOM

### 1. The canonical equations of motion; perturbation theory for nondegenerate systems

For several degrees of freedom ($f>1$) it rather suggests itself that we replace the representation of quantum-theoretical quantities by two-dimensional matrices by one in terms of $2f$-dimensional matrices, corresponding with the $2f$-dimensional manifold of stationary states in the classical $J$-space:

$$\begin{aligned} q_k &= \big(q_k(n_1 \ldots n_f, m_1 \ldots m_f)\big), \\ p_k &= \big(p_k(n_1 \ldots n_f, m_1 \ldots m_f)\big). \end{aligned} \tag{1}$$

Nevertheless this representation, albeit under certain circumstances

---

[1] Cf. the discussion at the end of § 4 of results obtained for $\nu_0=0$.

[2] Cf. the articles by W. Kuhn, Zs. f. Phys. **33** (1925) 408; W. Thomas, Naturwiss. **13** (1925) 627; F. Reiche and W. Thomas, Zs. f. Phys. **34** (1925) 510.

very convenient and clear, is by no means essential. Even for several degrees of freedom the fundamental dynamical equations assume the form of *matrix equations*, but these matrices can as heretofore also be written in two-dimensional form. It became apparent even for *one* degree of freedom that the *sequence* of the stationary states as given by the ordering of the matrix rows is (in contradistinction to the theory employed hitherto) purely fortuitous and is not governed by any intrinsic property of the system. This observation can now directly be referred to many-dimensional matrices too; one can carry out any arbitrary rearrangements and in particular transform the $2f$-dimensional matrices into two-dimensional ones. This is justified by the fact that the basic definitions of addition and multiplication, as also of differentiation with respect to time, are clearly independent of any *ordering relations* between the basis systems of indices $n_1, n_2, \ldots, n_f$, which taken *singly* specify the *states* and *in pairs* specify the *transitions*.

It is thence also clear that the general rules of matrix analysis, as presented in chapter 1 of Part I and in chapter 1 of the present paper, can be employed in the theory of systems having *several* degrees of freedom also. One can similarly take over the derivation of the equation of motion from the variational principle in I directly, so that we can in like manner write

$$\dot{q}_k = \frac{\partial H}{\partial p_k}; \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \tag{2}$$

The principal new feature distinguishable from those obtaining for systems with just one degree of freedom lies in the general commutation relations for $p_k$ and $q_k$ in the case of several degrees of freedom. Just as in the calculations for but one degree of freedom, so here also calculations with quantum-theoretical quantities would be to some extent indefinite if the 'commutation relations' were not specified.

As a plausible generalization of eqs. (5), ch. 1, the following equations suggest themselves:

$$\begin{aligned} p_k q_l - q_l p_k &= \frac{h}{2\pi i}\,\delta_{kl}, \\ p_k p_l - p_l p_k &= 0, \\ q_k q_l - q_l q_k &= 0, \end{aligned} \tag{3}$$

if $H$ denotes the (symmetrized) energy function, one can in consequence of these relations replace eqs. (2), ch. 2, by

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \tag{2'}$$

Further, it follows from these relations,[1] as in chapter 1 of the present paper, that

$$\begin{aligned} p_k f(q_1 \dots q_f, p_1 \dots p_f) - f p_k &= \frac{h}{2\pi i} \frac{\partial f}{\partial q_k}, \\ f q_k - q_k f &= \frac{h}{2\pi i} \frac{\partial f}{\partial p_k}. \end{aligned} \tag{4}$$

The proof of energy conservation and the frequency condition then follows from (2′) and (4), ch. 2, as shown in ch. 1. Similarly one can show with the aid of (3) and (4) that the canonical equations (2′), ch. 2, apply whenever the relations (3), ch. 2, are satisfied for a system $P_k$, $Q_k$ and the energy function is given as an analytical function of the $P_k$ and $Q_k$.

Thus a transformation of the variables $p_k$, $q_k$ into new variables $P_k$, $Q_k$ is termed 'canonical' if it leaves the relations (3), ch. 2, unaltered.

A very general class of such transformations is again given by the formulae

$$\begin{aligned} P_k &= S p_k S^{-1}, \\ Q_k &= S q_k S^{-1}. \end{aligned} \tag{5}$$

This transformation again has the property of converting every function $f(PQ)$ into

$$f(P_1, \dots, P_f, Q_1, \dots, Q_f) = S f(p_1, \dots, p_f, q_1, \dots, q_f) S^{-1}. \tag{6}$$

If a system $p_1^0, \dots, p_f^0, q_1^0, \dots, q_f^0$ is known, and satisfies the relations (3), ch. 2, then the problem of integrating eqs. (2), ch. 2, again reduces itself to the simpler problem: A function $S$ is to be sought, such that it satisfies the equations

$$\begin{aligned} p_k &= S p_k^0 S^{-1}, \\ q_k &= S q_k^0 S^{-1} \end{aligned} \tag{5a}$$

---

[1] The physical significance of these relations for dispersion theory is discussed by H. A. Kramers, Physika, December 1925.

and transforms $H$ into a diagonal matrix,

$$H(pq) = SH(p^0q^0)S^{-1} = W. \tag{7}$$

Equation (7) again represents the counterpart to the Hamilton partial differential equation.

Equations (3), ch. 2 would, together with (2), ch. 2, obviously entail too extensive a set of requirements for the $p_k$, $q_k$, if all these equations were *independent* of one another. As an interesting mathematical problem must rank the derivation of eqs. (3) using the least number of independent and mutually consistent assumptions; nevertheless, this question will not be handled here. We shall content ourselves with mentioning that

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_k (p_kq_k - q_kp_k) = 0$$

is a general outcome of the equations of motion (1), ch. 2. On the other hand, it will be shown generally that the eqs. (3), ch. 2, together with the equations of motion (2), ch. 2, or the equivalent requirement (7), ch. 2, *can* be satisfied (singular discrepancies apart, of course).

This proof is to be supplied in connection with the generalization of the perturbation theory presented in ch. 1 § 4, when extended to arbitrarily many degrees of freedom. We consider the energy function $H(pq)$ such that it can be written as

$$H = H_0(pq) + \lambda H_1(pq) + \lambda^2 H_2(pq) + \ldots, \tag{8}$$

so that

$$H_0(pq) = \sum_{k=1}^{f} H^{(k)}(p_kq_k).$$

Thus for $\lambda = 0$ we have $f$ *uncoupled* systems, each having a single degree of freedom; the $f$ cases

$$H = H^{(k)}(p_kq_k)$$

can be solved with

$$q_k = q_k^0, \qquad q_k = p_k^0,$$

wherein $q_k^0$, $p_k^0$ are *two-dimensional* matrices,

$$q_k^0 : (q_k^0(nm)); \qquad p_k^0 : (p_k^0(nm)). \tag{10}$$

If we formally regard these $f$ uncoupled systems as a single system having $f$ degrees of freedom, then $q_k^0, p_k^0$ would be represented as $2f$-dimensional matrices,

$$\left.\begin{aligned} q_k^0 &= \left(q_k^0(n_1 \dots n_f;\quad m_1 \dots m_f)\right), \\ p_k^0 &= \left(p_k^0(n_1 \dots n_f;\quad m_1 \dots m_f)\right), \end{aligned}\right\} \tag{11}$$

for which

$$\begin{aligned} q_k^0(n_1 \dots n_f;\quad m_1 \dots m_f) &= \delta_k q_k^0(n_k m_k), \\ p_k^0(n_1 \dots n_f;\quad m_1 \dots m_f) &= \delta_k p_k^0(n_k m_k), \end{aligned}$$

where $\delta_k = 1$ if $n_j = m_j$ for all $j$ except $j = k$, and $\delta_k = 0$ if for any $j (j \neq k)$, $n_j$ is not equal to $m_j$. Thence, however, one sees: firstly, that the equations

$$p_k^0 q_k^0 - q_k^0 p_k^0 = \frac{h}{2\pi i} \mathbf{1} \tag{12}$$

which originally obtained for the *two-dimensional* matrices (10), ch. 2, also hold for the *2f-dimensional* matrices (11), ch. 2; secondly, that the following relations ensue:

$$\begin{aligned} p_k^0 q_l^0 - q_l^0 p_k^0 &= 0 \quad \text{for} \quad l \neq k, \\ p_k^0 p_l^0 - p_l^0 p_k^0 &= q_k^0 q_l^0 - q_l^0 q_k^0 = 0. \end{aligned} \tag{13}$$

Hence for $\lambda = 0$, all the eqs. (13), ch. 2, indeed apply. It is to be shown that $p, q$ can be determined in such a manner that (3), ch. 2, is satisfied simultaneously with $H = W$ for the higher-order approximations also. One again assumes the system $H_0$ to have been chosen as *non-degenerate*, i.e., that on substituting $q = q^0$, $p = p^0$ no two diagonal elements of $H_0$ become identical. In this case we again have to set

$$q_k = S q_k^0 S^{-1}; \qquad p_k = S p_k^0 S^{-1} \tag{14}$$

as in eq. (5a), ch. 2, and to determine

$$S = \mathbf{1} + \lambda S_1 + \lambda^2 S_2 + \dots$$

in such a way as to satisfy the relation $H = W$. The eqs. (3), ch. 2, are then jointly also satisfied, since by virtue of (14) they go over into (12), (13). This completes the required proof.

Equations (3) are invariant with respect to a linear orthogonal transformation of the $q_k$ and $p_k$, for if one sets

$$q'_k = \sum_l a_{kl} q_l, \qquad \sum_l a_{kl} a_{jl} = \delta_{kj},$$
$$p'_k = \sum_l a_{kl} p_l,$$

then

$$p'_k q'_l - q'_l p'_k = \sum_{hj} a_{kh} a_{lj} (p_h q_j - q_j p_h) = \delta_{kl} \frac{h}{2\pi i}$$

and similarly for the other respective relations. If then the conditions (3), ch. 2, hold for a given Cartesian coordinate system, they will also be valid in every other Cartesian coordinate system.

By way of supplement, now that we have established (3), ch. 2, we demonstrate that a well-known law of classical mechanics is also compatible with the new theory.

Let

$$H = E_{\text{kin}} + E_{\text{pot}} = \tfrac{1}{2} \sum_k \frac{p_k^2}{m_k} + E_{\text{pot}}, \tag{15}$$

and let $E_{\text{pot}}$ be a homogeneous function of the coordinates of order $n$. Then from (3), ch. 2,

$$E_{\text{pot}} = \frac{1}{n} \sum_k \frac{\partial E_{\text{pot}}}{\partial q_k} q_k \tag{16}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_k p_k q_k = \sum_k (\dot{p}_k q_k + p_k \dot{q}_k) = 2E_{\text{kin}} - nE_{\text{pot}},$$

so that for the *mean* values,

$$\bar{E}_{\text{kin}} = \tfrac{1}{2} n \bar{E}_{\text{pot}}. \tag{17}$$

Hence, e.g., for $n=2$ (harmonic oscillations), $\bar{E}_{\text{kin}}=\bar{E}_{\text{pot}}$ and for $n=-1$ (Coulomb force), $\bar{E}_{\text{kin}}=-\tfrac{1}{2}\bar{E}_{\text{pot}}$.

## 2. Degenerate systems

We now turn to examination of degenerate systems. If we permit some of the frequencies $\nu(nm)$ to vanish (for simplicity, we imagine the matrices to be in two-dimensional representation), then *energy conservation*, $\dot{H}=0$ can still be derived from the considerations employed here and in Part I concerning the equations of motion and the commutation rules (3), ch. 2. But the relation $\dot{H}=0$ no longer necessarily implies that $H$ be a diagonal matrix and in consequence the proof of the frequency condition cannot be carried through. Thus for degenerate systems the equations of motion together with (3), ch. 2, do *not* alone suffice for the unique determination of the properties of a system: we need to strengthen these basic equations. An obvious assumption as to the form of this 'increase in rigour' is:

*For basic equations, one should be able generally to choose the commutation relations and the property*

$$H = W = \text{diagonal matrix}. \tag{18}$$

This requirement manifestly ensures the validity of the frequency condition for degenerate systems as well. Very probably, the energy $W$ is also thereby uniquely determined (apart from singular instances). On the other hand, the *coordinates* $q_k$ are *not* uniquely determined. Given a solution $p_k$, $q_k$ of $H(pq)=W$, we can get new solutions from

$$\begin{aligned} p' &= SpS^{-1}, \\ q' &= SqS^{-1}. \end{aligned} \tag{19}$$

Thence

$$H(p'q') = W' = SWS^{-1},$$

and the requirement $W'=W$ yields

$$WS - SW = \dot{S}\frac{h}{2\pi i} = 0,$$

and thus

$$S = \text{const.} \tag{20}$$

Let us at this stage examine this result as regards its implications for nondegenerate systems. From (2), ch. 2, the matrix $S$ has to become a diagonal matrix, and the eqs. (19), ch. 2, imply that

$$\begin{aligned} p'(nm) &= p(nm)S_n S_m^{-1}, \\ q'(nm) &= q(nm)S_n S_m^{-1}, \end{aligned} \tag{19'}$$

writing $S_n$ for $S(nn)$ for the sake of conciseness.

The uncertainty in the solution indicated hereby can significantly be reduced by the requirement that the new solution $p'$, $q'$ should also represent *'real'* motion, expressed in terms of Hermitian matrices, since this yields

$$|S_n S_n^{-1}| = |S_m S_m^{-1}|,$$

or

$$|S_n| = |S_m|. \tag{21}$$

Thus the indeterminacy which has here come to light represents an arbitrariness of the *phase constants*. We namely here find proof of the contention put forward in Part I that in each problem for every state $n$ a phase $\varphi_n$ always remains undetermined. From (19′) one can perceive the manner in which these phases enter into the elements of the matrices $\mathsf{p}$, $\mathsf{q}$. It was further conjectured in Part I that apart from the above-mentioned arbitrariness of phase for non-degenerate systems, no additional non-uniqueness is to be expected. It is clear that we could still add a constant matrix to each of the 'periodic' matrices $S_n$ in the perturbation calculations of ch. 1 § 4. However, this obviously does not imply that new phases which remain undetermined enter into each approximation. It is easy to see that utilization of this possibility cannot provide any more general solution $\mathsf{p}$, $\mathsf{q}$ provided that $\mathsf{p}^0$, $\mathsf{q}^0$ were right from the first taken to have undetermined phases.

If we now go over to degenerate systems, we cannot any longer infer from (20) that $\mathsf{S}$ is a diagonal matrix, and accordingly, using (19), we do indeed have the possibility of deriving solutions $\mathsf{p}'$, $\mathsf{q}'$ which are significantly different from $\mathsf{p}$, $\mathsf{q}$. This indeterminacy seems to lie in the very nature of things. Apparently, degenerate systems possess a *lability* by virtue of which arbitrarily small perturbations can bring about finite changes in coordinates, and this finds its mathematical expression in that in complete absence of perturbations, the solution of the dynamic equations remains partly indeterminate. Naturally, for every actual atom the coordinates which specify the physical properties of the system, in particular the transition probabilities, are always fixed uniquely either by external perturbations or by the previous history of the system.

Now we set out to examine the influence of arbitrary perturbations upon the degenerate system. We set

$$H(\mathsf{pq}) = H_0 + \lambda H_1 + \lambda^2 H_2 + \dots, \tag{22}$$

and let $\mathsf{p}^0$, $\mathsf{q}^0$ be an arbitrary, but definite, solution of the unperturbed problem:

$$H_0(p^0q^0) = W_0. \tag{23}$$

Then with

$$p = Sp^0S^{-1},$$
$$q = Sq^0S^{-1},$$

and with

$$S = S_0(\mathbf{1} + \lambda S_1 + \lambda^2 S_2 + \ldots), \tag{24}$$
$$S^{-1} = (\mathbf{1} - \lambda(S_1 + \lambda S_2 \ldots) + \lambda^2 \ldots)S_0^{-1}, \tag{25}$$

we find, on leaving out the arguments $p^0, q^0$ from $H_0, H_1,\ldots$:

$$S_0H_0S_0^{-1} = W_0, \tag{26}$$
$$S_0S_1H_0S_0^{-1} - S_0H_0S_1S_0^{-1} + S_0H_1S_0^{-1} = W_1, \tag{27}$$
$$S_0S_2H_0S_0^{-1} - S_0H_0S_2S_0^{-1} + S_0F_2(H_0H_1H_2; S_1)S_0^{-1} = W_2, \tag{28}$$

. . . . . . . . . . . . . . . . . . . . . . . .

$$S_0S_rH_0S_0^{-1} - S_0H_0S_rS_0^{-1}$$
$$+ S_0F_r(H_0H_1 \ldots H_r, S_1 \ldots S_{r-1})S_0^{-1} = W_r. \tag{29}$$

Thus we almost repeat eqs. (26), ch. 1, but with the difference that the left-hand sides are throughout multiplied on the left by $S_0$ and on the right by $S_0^{-1}$.

Equation (26), ch. 2, has already been cited above; $S_0(nm)$ becomes zero except for vanishing $\nu_0(nm)$. The remaining arbitrariness in $S_0$ now has to be used to advantage so far as possible in order to render the next equation soluble. Naturally, one cannot expect that every solution of $H{=}H_0$, and thus in particular the chosen solution $p^0, q^0$, will provide the limiting case $\lambda = 0$ of the solution $p, q$ of the problem (22), ch. 2. The function $S_0$ should serve to obtain from $p^0, q^0$ that solution of the degenerate problem which possesses this desired property.

We can rewrite eq. (27) as

$$S_1H_0 - H_0S_1 + H_1 = S_0^{-1}W_1S_0. \tag{30}$$

To make this soluble, one has to determine $S_0$ such that

$$\overline{H}_1 = S_0^{-1}W_1S_0 \tag{31}$$

for a *diagonal matrix* $W_1$. An indication as to how one can simultaneously satisfy this eq. (31) and the requirements dictated by (26), ch. 2, can here naturally just as little be given as that for the determination of secular perturbations in classical theory. We shall, however, later use a new algebraic method to arrive at a simple treatment of an extensive class of degeneracies (ch. 3).

If (31), ch. 2, is satisfied, (30), ch. 2, can be solved as in ch. 1. Thereby those terms $S_1(nm)$ of $S_1$ for which $\nu_0(nm)$ vanishes remain arbitrary, and this indeterminacy has to be utilized in order to solve the next higher order approximation formula, which can be transcribed as

$$S_2H_0 - H_0S_2 + F_2 = S_0^{-1}W_2S_0 \tag{32}$$

in order to fulfil the necessary relation

$$\overline{F_2(H_0, H_1, H_2; S_1)} = S_0^{-1}W_2S_0 \tag{31'}$$

with $W_2$ a diagonal matrix. This has to be satisfied for the problem to be soluble. The continuation of the procedure is clear.

The difficulty lies in the fact that at each order of approximation equations have to be satisfied by matrices which are already fixed to a large extent, so that it is not perceptible whether or not these equations will really prove soluble. In classical theory there is, though, an altogether analogous difficulty. These difficulties can, at least in the higher orders of approximation, be removed if in some approximation the system becomes nondegenerate.

Suppose, for example, that $p^{(1)}$ and $q^{(1)}$ in

$$q = q^0 + \lambda q^{(1)} + \dots,$$
$$p = p^0 + \lambda p^{(1)} + \dots$$

have really been determined, so that with

$$Q = q_0 + \lambda q^{(1)}$$
$$P = p^0 + \lambda p^{(1)}$$

one has

$$H(PQ) = W_0 + \lambda W_1 + \lambda^2 H_2' + \lambda^3 H_3' + \dots,$$

and suppose

$$\nu_0(nm) + \lambda\nu_1(nm) \neq 0 \quad \text{for} \quad n \neq m.$$

If for brevity we write $H'_0$ for $W_0+\lambda W_1$ and set

$$p = SPS^{-1}$$
$$q = SQS^{-1},$$

then we have to build the following relation,

$$S(H'_0 + \lambda^2 H'_2 + \lambda^3 H'_3 + \ldots)S^{-1} = W,$$

which, with the procedures of ch. 1, can be achieved with

$$S = \mathbf{1} + \lambda^2 S_2 + \lambda^3 S_3 + \ldots.$$

The generalization of these considerations for the case in which only in the $r$th approximation can one attain a nondegenerate system $W=W_0+\lambda W_1+\ldots+\lambda^r W_r$ follows of itself.[1]

In conclusion, we deem it important to point out that the notorious convergence difficulties encountered in the classical perturbation series, which play so decisive a rôle in the discussion of the three-body problem, do not arise here in quantum-mechanical perturbation theory; rather, one would here in general expect finite orbits to be periodic also.

## CHAPTER 3. CONNECTION WITH THE THEORY OF EIGENVALUES OF HERMITIAN FORMS

### 1. General method

The treatment in the preceding sections has aimed at solving the basic quantum-theoretical equations in a manner as closely parallel to classical theory as possible. But behind the formalism of this perturbation theory there lurks a very simple, purely algebraic connection and it is well worth while to bring this into the limelight. Apart from the deeper insight into the mathematical structure of the theory, we thereby gain the advantage of being able to use the methods and results developed earlier in mathematics. We shall thus arrive at a new definition of the energy constants ('terms') which remains valid in the case of aperiodic motion also, i.e., of continuously-varying indices. Thereby we attain the prospect of finding methods for direct

---

[1] Analogous cases in classical mechanics have been discussed by M. Born and W. Heisenberg, Ann. d. Phys. **74** (1924) 1.

calculation of the energy without explicitly solving the problem of motion: methods which correspond to Sommerfeld's method of complex integration. We shall then be able to treat perturbations of an extensive class of degenerate systems completely, which the above-mentioned perturbation methods were not yet able to handle.

In considering a problem of $f$ degrees of freedom specified by the energy function $H(pq)$, we can first select any system of matrices $p_k^0$, $q_k^0$ whatsoever such that at all events the commutation relations (3), ch. 2, are satisfied: for example, we can take the $p_k$, $q_k$ for a system of noncoupled harmonic oscillators.

Then, as mentioned in ch. 2 § 1, the dynamic problem, e.g., the determination of the $p_k$, $q_k$ can be formulated as: A transformation $(p_k^0 q_k^0) \to (p_k q_k)$ is to be found which leaves eqs. (3), ch. 2, invariant and at the same time reduces the energy to a diagonal matrix.

The transformation of matrices can most easily be grasped if one regards them as a system of coefficients for linear transformations or bilinear forms. We therefore premise some known results of the algebra of such forms.

To every matrix $\mathsf{a} = (a(nm))$ there belongs a *bilinear form*

$$A(xy) = \sum_{nm} a(nm) x_n y_m \tag{1}$$

of two series of variables $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$. If the matrix be Hermitian, i.e., if the *transposed matrix* $\tilde{\mathsf{a}} = (a(mn))$ be equal to the complex conjugate of the original matrix,

$$\tilde{\mathsf{a}} = \mathsf{a}^*, \qquad a(mn) = a^*(nm), \tag{2}$$

then the form $A$ assumes real values if in place of the variables $y_n$ one substitutes the complex conjugate values $x_n$:

$$A(xx^*) = \sum_{nm} a(nm) x_n x_m^*. \tag{1a}$$

We recall the readily demonstrable transposition rule

$$(\widetilde{\mathsf{ab}}) = \tilde{\mathsf{b}}\,\tilde{\mathsf{a}} \tag{3}$$

and now subject the $x_n$ to a linear transformation

$$x_n = \sum_{l} v(ln) y_l \tag{4}$$

with the aid of the (complex) matrix $\mathsf{v} = (v(ln))$.

Then the form $A$ goes over into

$$A(xx^*) = B(yy^*) = \sum_{nm} b(nm) y_n y_m^*, \tag{5}$$

with

$$b(nm) = \sum_{kl} v(nk) a(kl) v^*(ml),$$

or, in matrix notation,

$$\boldsymbol{b} = \boldsymbol{v a v^*}. \tag{6}$$

This is termed the generation of a matrix $\boldsymbol{b}$ by the transformation $\boldsymbol{v}$ applied to $\boldsymbol{a}$.

The matrix $\boldsymbol{b}$ is again of Hermitian type, for, with (3), ch. 3,

$$\tilde{\boldsymbol{b}} = \boldsymbol{v}^* \tilde{\boldsymbol{a}} \tilde{\boldsymbol{v}} = \boldsymbol{v}^* \boldsymbol{a}^* \tilde{\boldsymbol{v}} = \boldsymbol{b}^*. \tag{7}$$

The matrix $\boldsymbol{v}$ is called *orthogonal* if the respective transformation leaves the Hermitian unit form

$$E(xx^*) = \sum_{n} x_n x_n^*$$

invariant; from the result derived above, this is the case if and only if

$$\boldsymbol{v} \tilde{\boldsymbol{v}}^* = \mathbf{1}, \quad \text{or} \quad \tilde{\boldsymbol{v}}^* = \boldsymbol{v}^{-1}. \tag{8}$$

Thus, for instance, the permutation matrices mentioned in ch. 1 § 2 are real orthogonal matrices.

As is known, it is always possible for a finite number of variables to effect an orthogonal transformation of a form into a sum of squares (transformation to principal axes).[1]

$$A(xx^*) = \sum_{n} W_n y_n y_n^*. \tag{9}$$

For matrices, this means: a matrix exists for which

$$\boldsymbol{v} \tilde{\boldsymbol{v}}^* = \mathbf{1} \quad \text{and} \quad \boldsymbol{v a} \tilde{\boldsymbol{v}}^* = \boldsymbol{v a v}^{-1} = \boldsymbol{W}, \tag{10}$$

where $\boldsymbol{W} = (W_n \delta_{nm})$ is a diagonal matrix.

For infinite matrices, all the cases investigated so far have been found to obey an analogous rule; it can however occur that the index $n$ on the right-hand side runs not only through a set of discrete numbers

---

[1] We write the coefficients of the transformed form $W_n$ because in quantum mechanics they stand for the 'energy'.

but also through a continuous range of values; this would correspond[1] to an integral constituent of (9) and the transformation (4).

The quantities $W_n$ are termed 'eigenvalues', their ensemble is the 'mathematical spectrum' of the form, made up of 'point-' and 'continuous' spectrum. As we shall see, this is identical with the 'term-spectrum' in physics, whereas the 'frequency spectrum' is obtained from this by forming differences.

This transformation to principal axes now directly presents us with the solution of our dynamic problem which consists in seeking a transformation $(p^0q^0) \to (pq)$ such that the eqs. (3), ch. 2 are left invariant and at the same time the energy is brought into diagonal matrix form.

By the above rules of algebra, there exists an orthogonal matrix $S$ for which

$$S\tilde{S}^* = \mathbf{1}, \qquad \tilde{S}^*S = \mathbf{1} \tag{11}$$

and for which the transformations

$$\begin{aligned} p_k &= Sp_k^0\tilde{S}^* = Sp_k^0S^{-1}, \\ q_k &= Sq_k^0\tilde{S}^* = Sq_k^0S^{-1} \end{aligned} \tag{12}$$

leave

(i) the Hermitian character of $p_k^0$, $q_k^0$ conserved also for the $p_k$, $q_k$;
(ii) the eqs. (3), ch. 2, invariant;
(iii) the energy

$$H(pq) = SH(p^0q^0)S^{-1} = W \tag{13}$$

converted into diagonal matrix form.

We wish to discuss the question of the uniqueness of this solution and in particular whether one could not generate other energy values through another orthogonal transformation $T$. Let us assume that $W'$, as given by

$$TH(p^0q^0)T^{-1} = W',$$

is a diagonal matrix which differs from $W$. One would then have

$$TS^{-1}SHS^{-1}ST^{-1} = TS^{-1}W(TS^{-1})^{-1},$$

---

[1] Up till now, the theory of quadratic (or Hermitian) forms of infinitely many variables has been developed mainly for a special class ('bounded' forms) (D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen; E. Hellinger, Crelles Journ. **136** (1910) 1). But here we are concerned just with non-bounded forms. We may nevertheless assume that in the main the rules run likewise.

and our question is equivalent to asking whether it is possible, starting from a diagonal matrix $W$ to build another, $W'$, through te transformation

$$W' = MWM^{-1}, \qquad M\tilde{M}^* = 1 \tag{14}$$

such that $W'$ can *not* be derived from $W$ by a permutation of the diagonal elements.

However, eq. (14), ch. 3, can be written

$$W'M - MW = 0.$$

and thus implies

$$M(nm)(W'_n - W_m) = 0. \tag{14a}$$

From the orthogonality of $M$, it follows in particular for $m=n$ that

$$\sum_k |M(nk)|^2 = 1, \qquad \sum_k |M(kn)|^2 = 1;$$

and consequently for a fixed $n$ neither all the $M(nk)$ nor all the $M(kn)$ can vanish. But then (14a), ch. 3, shows that for every $n$ there is certainly an $m$ for which $W'_n = W_m$, i.e., all the $W'_n$ appear among the $W_m$. The same holds inversely.

Thus all solutions derived from (12), ch. 3, lead (for given $p_k^0$, $q_k^0$) to the same values for the energies of the stationary states, in accord with the conjecture stated in ch. 2 that the energies are always uniquely determined by the fundamental dynamic equations.

Degenerate systems will be characterized by the fact that multiple eigenvalues occur. The multiplicity of the eigenvalue $W_n$, i.e., the number of linear independent solutions $v(ln)$ of eq. (4), ch. 3, yields the statistical weight of the respective state.

The importance of eq. (9), ch. 3, for our physical theory lies in the fact that various methods[1] exist in the algebra of finite or bounded infinite forms for determining the eigenvalues of a form without actually carrying the transformation through. It is to be hoped that such methods will prove of much avail in the future treatment of certain physical systems.

---

[1] For finite forms, the eigenvalues are the roots of an algebraic equation. Here, and also for bounded infinite matrices, they can be determined, e.g., by the method of Graeffe and Bernoulli; see, for example, R. Courant and D. Hilbert, Methoden der mathematischen Physik **1** (Springer, Berlin, 1924) § 3, pp. 14, 15.

## 2. Application to perturbation theory

In the following, we show that our present algebraic conception of the dynamic problem not only leads to exactly those formulae which were previously derived in ch. 1 § 4 in connection with perturbation theory in classical mechanics, but that when applied to degenerate systems it is considerably superior to the theory used hitherto.

We thus again assume that $H$ has the form

$$H = H_0 + \lambda H_1 + \lambda^2 H_2 + \dots,$$

and that the dynamic problem specified by $H_0$ has the solution $p_k^0$, $q_k^0$. We take these quantities as our starting coordinates from which the $p_k$, $q_k$ are to be found, using an orthogonal transformation $S$. Naturally, the form assumed for $H$ does not basically represent any limitation in generality, inasmuch as one can obviously separate off from $H$ a component $H_0$ of any desired form; however, the convergence of the power series in $\lambda$ will depend essentially upon an apposite choice of $H_0$.

To undertake a principal-axes transformation of the Hermitian form

$$\sum_{mn} H_{mn} x_m x_n^*$$

we can, as is known, proceed as follows:
We attempt to find a solution of the linear equations

$$W x_k - \sum_l H(kl) x_l = 0; \tag{15}$$

this is possible only for certain values of the parameter $W$, namely $W = W_n$, when $W_n$ again denotes the eigenvalues (energy values). We first assume that no degeneracy is present, so that all $W_n$ are different. Then to each $W_n$ there corresponds a solution $x_k = x_{kn}$ (determined except for a multiplicative factor), and hence the identities

$$W_n x_{kn} - \sum_l H(kl) x_{ln} = 0,$$

$$W_m x_{km}^* - \sum_l H^*(kl) x_{lm}^* = 0$$

obtain. On multiplying the former by $x_{km}^*$, the latter by $x_{kn}$ and summing over $k$, it follows on subtraction (because of the Hermitian character of $H$) that

$$(W_n - W_m) \sum_k x_{kn} x_{km}^* = 0.$$

By choosing the proportionality factor suitably, one can normalize to

$$\sum_k x_{kn} x_{kn}^* = 1.$$

Hence the $x_{kn}$ form an orthogonal matrix

$$S = (x_{kn}).$$

It is precisely this which transforms the given form to a sum of squares, since if we substitute

$$x_k = \sum_n x_{kn} y_n$$

into the form, we obtain

$$\begin{aligned} \sum_{kl} H(kl) x_k x_l^* &= \sum_{kl} \sum_{mn} H(kl) x_{km} x_{ln}^* y_m y_n^* \\ &= \sum_{mn} \sum_{l} W_m x_{lm} x_{ln}^* y_m y_n^* \\ &= \sum_m W_m y_m y_m^*. \end{aligned}$$

From our assumption as to the form of $H$, the coefficients of eq. (15), ch. 3, now have the form

$$H(kl) = \delta_{kl} W_l^0 + \lambda H_1(kl) + \lambda^2 H_2(kl) + \ldots.$$

We thus seek to find the solution of (15), ch. 3, through expansions of the form

$$\begin{aligned} W &= W^0 + \lambda W^{(1)} + \lambda^2 W^{(2)} + \ldots \\ x_k &= x_k^0 + \lambda x_k^{(1)} + \lambda^2 x_k^{(2)} + \ldots. \end{aligned} \tag{16}$$

If we substitute the above in (15), ch. 3, we obtain the approximation equations

$$\begin{aligned} &\text{(a)} \quad x_k^0 (W^0 - W_k^0) = 0, \\ &\text{(b)} \quad x_k^{(1)} (W^0 - W_k^0) = -x_k^0 W^{(1)} + \sum_l H^{(1)}(kl) x_l^0, \\ &\text{(c)} \quad x_k^{(2)} (W^0 - W_k^0) = -(x_k^{(1)} W^{(1)} + x_k^{(0)} W^{(2)}) \\ &\qquad\qquad + \sum_l \left( H^{(1)}(kl) x_l^{(1)} + H^{(2)}(kl) x_l^0 \right). \end{aligned} \tag{17}$$

It follows from (17a), ch. 3, that $W$ has to become equal to one of the $W_k$, since otherwise all $x_k^0$ would vanish and we could then also infer the vanishing of $x_k^{(1)}, x_k^{(2)}, \ldots$ in sequence from the subsequent approximation equations.

If, then, we take our starting system as nondegenerate, and thus all the $W_k^0$ as different from one another, the solution of (17a), ch. 3, is

$$W = W_n^0; \qquad x_{nn}^0 = y_n^0; \qquad x_{kn}^0 = 0 \qquad \text{for} \qquad k \neq n. \tag{18}$$

Herein, $y_n^0$ is an arbitrary number.

If we substitute this in (17b), ch. 3, we find, depending upon whether $k=n$ or $k \neq n$,

$$\begin{aligned} 0 &= y_n^0\left(-W^{(1)} + H^{(1)}(nn)\right), \\ x_k^{(1)}(W_n^0 - W_k^0) &= H^{(1)}(kn) y_n^0, \qquad k \neq n. \end{aligned}$$

Thus the solution runs

$$\begin{gathered} W^{(1)} = H^{(1)}(nn); \qquad x_{nn}^{(1)} = y_n^{(1)}; \\ x_{kn}^{(1)} = -\frac{H^{(1)}(kn)}{h\nu_0(kn)} y_n^0 \qquad \text{for} \qquad k \neq n, \end{gathered} \tag{19}$$

where again $y_n^{(1)}$ is an arbitrary number.

Hence it similarly follows from (17c), ch. 3, that

$$\begin{aligned} W^{(2)} &= H^{(2)}(nn) - \frac{1}{h} \sum_l{}' \frac{H^{(1)}(nl) H^{(1)}(ln)}{\nu_0(ln)}, \\ x_{nn}^{(2)} &= y_n^{(2)} \\ x_{kn}^{(2)} &= \left( \frac{1}{h^2} \sum_l{}' \frac{H^{(1)}(kl) H^{(1)}(ln)}{\nu_0(kn)\nu_0(ln)} - \frac{H^{(1)}(nn) H^{(1)}(kn)}{h^2 \nu_0(kn)^2} \right. \\ &\qquad \left. - \frac{H^{(2)}(kn)}{h\nu_0(kn)} \right) y_n^0 - \frac{H^{(1)}(kn)}{h\nu_0(kn)} y_n^{(1)}. \end{aligned} \tag{20}$$

The solution of the third-order approximation can be derived just as easily; we cite only the energy value:

$$\begin{aligned} W^{(3)} &= H^{(3)}(nn) - \frac{1}{h} \sum_l{}' \frac{H^{(1)}(nl) H^{(2)}(ln) + H^{(2)}(nl) H^{(1)}(ln)}{\nu_0(ln)} \\ &+ \frac{1}{h^2} \left( \sum_{kl}{}' \frac{H^{(1)}(nl) H^{(1)}(lk) H^{(1)}(kn)}{\nu_0(ln)\nu_0(kn)} - H^{(1)}(nn) \sum_l{}' \frac{H^{(1)}(nl) H^{(1)}(ln)}{\nu_0(ln)^2} \right). \end{aligned}$$

The quantities $y_n^{(0)}, y_n^{(1)}, \ldots$, which for the present are arbitrary, serve to normalize the solution (it is orthogonal of itself); the condition

$$\sum_k x_{kn} x_{kn}^* = 1$$

yields, for

$$x_{kn} = x_{kn}^0 + \lambda x_{kn}^{(1)} + \lambda^2 x_{kn}^{(2)} + \dots,$$

the equations

$$\sum_k x_{kn}^0 x_{kn}^{*0} = 1$$

$$\sum_k (x_{kn}^0 x_{kn}^{*(1)} + x_{kn}^{(1)} x_{kn}^{*0}) = 0$$

$$\dots\dots\dots\dots$$

On substituting the solution just obtained, it follows successively that

$$|y_n^0|^2 = 1$$

$$y_n^0 y_n^{*(1)} + y_n^{*0} y_n^{(1)} = 0$$

$$\dots\dots\dots$$

If we now set

$$y_n^{(p)} = a_n^{(p)} e^{i\varphi_n(p)}, \tag{21}$$

we obtain

$$a_n^0 = 1$$

$$2a_n^{(1)} \cos(\varphi_n^0 - \varphi_n^{(1)}) = 0$$

$$\dots\dots\dots\dots\dots\dots$$

$$2a_n^{(r)} \cos(\varphi_n^0 - \varphi_n^{(r)}) = F^{(r)}(a^{(r-1)}, \varphi^{(r-1)}, \dots).$$

Thus the phase constants $\varphi_n^0, \varphi_n^{(1)}, \dots$ can be chosen arbitrarily; the $a_n^0, a_n^{(1)}, \dots$ can be evaluated in sequence and determined uniquely. This stands in agreement with the result we found earlier (§ 3), namely that the phases of the diagonal terms of S remain undetermined.

On substituting the values $a_n^0 = 1, \dots$ obtained above into (21), ch. 3, and this in turn into (18), (19), (20), ch. 3, we see that the 'perturbation procedure' carried through earlier yielded just the solution for which the phases $\varphi_n^{(p)}$ vanish, i.e., for which the diagonal terms of S are real.

We now turn to consideration of the case in which the starting system is degenerate and in which $W_n^0$ is an $r$-fold eigenvalue. This means that eq. (17a), ch. 3, has the solution

$$\begin{gathered} W = W_n; \qquad x_{nn}^0 = y_{1,n}^0, \qquad x_{n,n+1}^0 = y_{2,n} \qquad \dots \\ x_{n,n+r-1} = y_{r,n}, \\ x_{kn}^0 = 0 \qquad \text{for} \qquad k \neq n, n+1, \dots, n+r-1. \end{gathered} \tag{23}$$

The left-hand side of (17b), ch. 3, then vanishes for

$$k = n, n+1, \ldots, n+r-1;$$

this yields ($r$) equations:

$$W^{(1)} y^0_{kn} - \sum_{l=1}^{r} H^{(1)}(n+k, n+l) y^0_{ln} = 0; \qquad k = 1, 2, \ldots, r, \tag{24}$$

whose array of coefficients is again of Hermitian type.

On setting the determinant to zero, one obtains a secular equation of the $r$th order for $W^{(1)}$:

$$\det\left(W^{(1)}\delta_{kl} - H^{(1)}(n+k, n+l)\right) = 0, \tag{25}$$

whose roots are certainly real. To each root there belong one or more independent solutions of eqs. (24), ch. 3.

If one selects one of these solutions, the perturbation procedure can be pursued: we shall, however, not go into this further here.

It suffices to have recognized that our algebraic method is able to handle all degeneracies of finite multiplicity, i.e., that it can reduce the problem to the solution of algebraic equations. If, for example, each eigenvalue occurs twice, so that to each there belongs a vanishing frequency $\nu_0(nm)$, the perturbation problem leads to a quadratic equation:

$$\begin{vmatrix} W^{(1)} - H^{(1)}(n, n) & -H^{(1)}(n, n+1) \\ -H^{(1)}(n+1, n) & W^{(1)} - H^{(1)}(n+1, n+1) \end{vmatrix} = 0.$$

This case obtains when two originally identical nondegenerate systems (in which all frequencies in each of the respective systems are to be different) are coupled through some force.

Further, the orthogonality relation

$$\sum_k x^0_{kn} x^{*0}_{kn} = 1$$

has an interesting meaning in the case of degenerate systems. Because of (23), this relation goes over into

$$\sum_{l=1}^{r} y^0_{ln} y^{*0}_{ln} = 1.$$

From this it follows that, if $m$ denotes any number in the series $n, n+1, \ldots, n+r-1$, and $k$ denotes any number outside this set, the sums

$$\sum_{m=n}^{n+r-1} p^0(mk)p^{*0}(mk),$$

$$\sum_{m=n}^{n+r-1} q^0(mk)q^{*0}(mk)$$

are uniquely determined, even for degenerate systems, e.g., the summations are invariant with respect to those transformations which, by (19), ch. 2, allow new and altogether different solutions $p'$, $q'$ to arise from certain solutions $p$, $q$ in the case of degeneracy. This result provides a mathematical representation of the so-called spectroscopic stability, which has played an important part in the more recent theories of fine-structure intensities (cf. ch. 4).

## 3. Continuous spectra

The simultaneous appearance of both continuous and line spectra as solutions of the same equations of motion and the same commutation relations seemed to us to represent a particularly significant feature of the new theory. In spite of this close connection between the two kinds of spectra, there nevertheless are characteristic distinctions, both mathematically and physically, between continuous and discrete spectra, corresponding to the difference between Fourier series and Fourier integrals in classical theory; it therefore strikes us as desirable to indicate the rough outlines of the treatment of continuous spectra here. The mathematical theory of continuous spectra which occur for infinite quadratic forms has, starting from the fundamental investigations of Hilbert, explicitly been developed by Hellinger (*loc.cit.*) for the case of bounded quadratic forms. If we here permit ourselves to take over Hellinger's results to the unbounded forms which appear in our case, we feel ourselves to be justified by the fact that Hellinger's methods obviously conform exactly to the physical content of the problem posed.

Let us first briefly examine the classical analogue to our problem, namely aperiodic motion and its Fourier integral. Whereas in a Fourier series a certain amplitude $a(\nu)$ always belongs to an oscillation $\exp(2\pi i\nu t)$, in the case of a Fourier integral one has a quantity of the form $\varphi(\nu)\mathrm{d}\nu$ in place of $a(\nu)$, where $\varphi(\nu)$ might in a sense be conceived as an amplitude-density per frequency interval $\mathrm{d}\nu$. In a similar and physically immediately obvious manner, one can always relate all quantities such as intensity, polarization, etc. to a frequency interval $\mathrm{d}\nu$ between $\nu$ and $\nu+\mathrm{d}\nu$, but never to a definite frequency itself. We shall have to expect quite similar conditions to apply in quantum mechanics. Instead of quantities $q(kl)$ we shall have quantities of the form $q(k, W)\mathrm{d}W$ or $q(W, W')\mathrm{d}W\mathrm{d}W'$, depending upon whether one or both of the two indices lie in the continuous region. Indeed, in place of the energy $W$ itself, there will have to be a 'total energy' per interval $\mathrm{d}W$, since the probability for an atom to have an absolutely definite energy $W$ in the continuous region is zero. To elucidate these questions we shall in the following briefly sketch Hellinger's mathematical theory.

For infinite quadratic forms, the case may arise that the form

$$\sum_{mn} H(mn)x_m x_n^*$$

cannot be converted into the expression $\sum_n W_n y_n y_n^*$ by an orthogonal substitution. We may then assume, in analogy with the results for bounded forms, that a representation with a continuous spectrum exists,

$$\sum_{mn} H(mn)x_m x_n^* = \sum_n W_n y_n y_n^* + \int W(\varphi)y(\varphi)y^*(\varphi)\mathrm{d}\varphi, \tag{26}$$

in which the original variables are connected with new variables $y_n$, $y(\varphi)$ through an 'orthogonal transformation'; one only has to specify more clearly what is here understood by an orthogonal transformation.

If we again consider the linear equations (15), ch. 3,

$$Wx_k - \sum_l H(kl)x_l = 0, \tag{27}$$

the case under review in which (26), ch. 3, contains an integral component will occur when there are not only discrete values $W_n$, for which these equations can be solved, but also a continuum of such values comprising one or more 'segments' on the $W$-axis (continuous spectrum). For any given point $W$ of this continuum, there exists a solution $x_l(W)$ (or several, which we for simplicity wish to exclude); for two such $W$-values, $W'$ and $W''$, the equations

$$\begin{aligned} W'x_k(W') - \sum_l H(kl)x_l(W') &= 0, \\ W''x_k^*(W'') - \sum H^*(kl)x_l^*(W'') &= 0 \end{aligned} \tag{28}$$

obtain, from which, as above, we conclude that

$$(W' - W'') \sum_k x_k(W') x_k(W'') = 0. \tag{29}$$

If one tries imposing the normalization condition

$$\sum_k |x_k(W)|^2 = 1$$

on top of these orthogonality relations, one observes that the function of two variables

$$\sum_k x_k(W') x_k(W'')$$

becomes wildly irregular, if it exists at all. The above sum does not in fact converge and therefore does not represent a function.

Accordingly, a different type of normalization is required. With Hellinger, we set

$$\sum_k |\int x_k(W) \mathrm{d}W|^2 = \varphi(W). \tag{30}$$

The series on the left-hand side is in general convergent and represents a monotonous function $\varphi(W)$, which apart from certain restrictions can be chosen arbitrarily, since the $x_k(W)$ are of course determined only up to a factor which is independent of $k$. We shall later discuss the physical significance of this function $\varphi(W)$, by which the solutions $x_k(W)$ are defined. Hellinger has termed $\varphi(W)$ the 'basis function' and has shown that the orthogonality conditions can be derived in the following form: If $\Delta_1$ and $\Delta_2$ be any two intervals of the continuous spectrum and $\Delta_{12}$ the interval common to them both (which may also be absent), then

$$\begin{aligned} \sum_k \int_{\Delta_1} x_k(W')\,\mathrm{d}W' \int_{\Delta_2} x_k(W'')\,\mathrm{d}W'' &= \int_{\Delta_{12}} \mathrm{d}\varphi(W) \\ &= \varphi(W^{(2)}) - \varphi(W^{(1)}), \end{aligned} \tag{31}$$

where $W^{(1)}$, $W^{(2)}$ are the end-points of $\Delta_{12}$. Hence if there is no overlap between the intervals $\Delta_1$, $\Delta_2$, a zero stands on the right-hand side.

If one conceives the intervals $\Delta_1$, $\Delta_2$, $\Delta_{12}$ to be very small, one can symbolically write

$$\sum_k x_k(W')\,\mathrm{d}W' \cdot x_k(W'')\,\mathrm{d}W'' = \mathrm{d}\varphi(W). \tag{32}$$

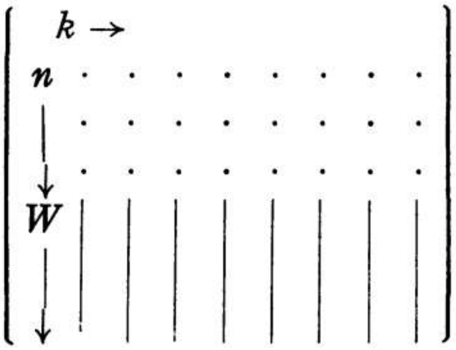
This relation prompts the suggestion that one operate generally with the quantities $x_k(W)\,\mathrm{d}W$ as 'differential solutions' of (27), ch. 3, whereby one has to note that the respective equations are always to be interpreted in the sense of (31), ch. 3. These differential solutions are orthogonal in the usual way, but instead of being normalized to unity, are normalized to the differential of the basis function $\varphi(W)$.

The totality of discrete values $x_{kn}$, and of values $x_k(W)$ which are discrete in one index and have a continuous distribution in the other, comprises the elements of the 'orthogonal' matrix

$$S = \left(x_{kn}, x_k(W)\,\mathrm{d}W\right),$$

which can schematically be represented as:

$$S =$$



$$. \qquad (33)$$

The orthogonality and normalization equations for the entire matrix split into four different groups:

$$\sum_k x_{km}x^*_{kn} = \delta_{mn};$$
$$\sum_k x_{kn}x^*_k(W)\,\mathrm{d}W = 0; \qquad \sum_k x_k(W)\,\mathrm{d}W\cdot x^*_{kn} = 0; \qquad (34)$$
$$\sum_k x_k(W')\,\mathrm{d}W'\cdot x^*_k(W'')\,\mathrm{d}W'' = \mathrm{d}\varphi.$$

We can also write the orthogonality relations for the columns, which read

$$\sum_n x_{kn}x^*_{ln} + \int \frac{x_k(W)\,\mathrm{d}W\cdot x^*_l(W)\,\mathrm{d}W}{\mathrm{d}\varphi}$$
$$= \sum_n x_{kn}x^*_{ln} + \int \frac{\mathrm{d}W}{\varphi'}\,x_k(W)x^*_l(W) = \delta_{kl}, \qquad (35)$$

where the prime denotes differentiation, $\varphi' = \mathrm{d}\varphi/\mathrm{d}W$.

With the aid of this matrix, we have to transform the variables $x_k$ into new ones, $y_n$, $y(\varphi)\,\mathrm{d}\varphi$. We set:

$$y_n = \sum_k x_{kn}\cdot x_k,$$
$$y(\varphi)\,\mathrm{d}\varphi = \sum_k x_k(W)\,\mathrm{d}W\cdot x_k. \qquad (36)$$

A simple calculation then yields

$$\sum_n W_n y_n y_n^* + \int W(\varphi) y(\varphi) y^*(\varphi)\, \mathrm{d}\varphi = \sum_{kl} H(kl) x_k x_l^*. \tag{37}$$

The principal-axes transformation has thereby been carried through.

Let us now investigate which representation of coordinate and momentum matrices is obtained with the aid of this orthogonal transformation, e.g., what is meant here by the equations

$$\begin{aligned} \mathbf{p} &= \mathsf{S}\mathbf{p}_0\mathsf{S}^{-1}, \\ \mathbf{q} &= \mathsf{S}\mathbf{q}_0\mathsf{S}^{-1}, \end{aligned} \tag{38}$$

or, generally, by

$$\mathbf{f}(\mathbf{pq}) = \mathsf{S}\mathbf{f}(\mathbf{p}_0\mathbf{q}_0)\mathsf{S}^{-1}. \tag{39}$$

We find, for example, four types of elements for $\mathbf{p}$:

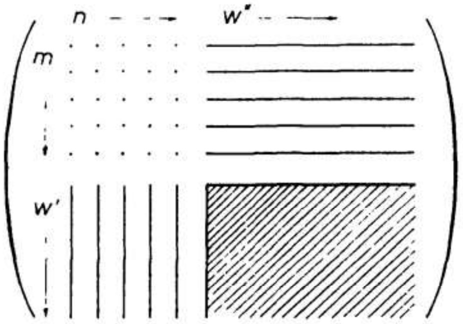
$$\begin{aligned} p(mn) &= \sum_{kl} x_{km}^* p^0(kl) x_{ln} \\ p(m, W)\, \mathrm{d}W &= \sum_{kl} x_{km}^* p^0(kl) x_l(W)\, \mathrm{d}W, \\ p(W, n)\, \mathrm{d}W &= \sum_{kl} x_k^*(W)\, \mathrm{d}W \cdot p^0(kl) x_{ln}, \\ p(W', W'')\, \mathrm{d}W'\, \mathrm{d}W'' &= \sum_{kl} x_k^*(W')\, \mathrm{d}W' p^0(kl) x_l(W'')\, \mathrm{d}W''. \end{aligned} \tag{40}$$

In a similar manner, instead of the amplitudes $p(mn)$, 'amplitude densities' $p(mW)\, \mathrm{d}W$ (which refer to an interval $\mathrm{d}W$) occur generally in the case of a continuously variable index. This accords with our previously declared expectation. It is, however, not necessary to take just the energy as the continuously variable index. In place of the energy, one could, for example, introduce the quantity $\varphi(W)$. Then in place of $p(mW)\, \mathrm{d}W$ one would have $p(m\varphi)(\mathrm{d}W/\mathrm{d}\varphi)\, \mathrm{d}\varphi$. Finally, in the continuous case the energy $W_n$ is replaced by the quantity $W(\varphi)\, \mathrm{d}\varphi$. In place of the energy of the individual atom, we get a sort of total energy per interval $\mathrm{d}W$. Thence $\mathrm{d}\varphi$ essentially represents the number of atoms having an energy which lies between $W$ and $W+\mathrm{d}W$, or the *a priori* probability that the energy of the atom lies between $W$ and $W+\mathrm{d}W$. We here most clearly observe the difference between the cases with discrete stationary states on the one hand and those with a continuous manifold of states on the other hand, and we can see a simple connection between the problem of statistical weights and the question of the normalization of the solution of (27), ch. 3. In the case of discrete states when there are no multiple eigenvalues, we make the simple physical contention that each state should have the statistical weight 1. This was ensured by the fact that we normalized the $x_{kn}$ on the basis of the requirement

$$\sum_k x_{kn}x_{kn}^* = 1.$$

In the case of continuous manifolds of states, it was not possible to fix the *a priori* probabilities so simply; more detailed investigations of the problem in question are necessary for their determination and hence also for the evaluation of the function $\varphi$. Hence the connection between transition probabilities and the amplitudes might also assume a somewhat more complicated aspect in the case of continuous spectra than for line spectra.

The matrices of $p$, $q$ or $f(p, q)$ represented by (40), ch. 3, and corresponding forms, can for the general case be made clear by the adjoining scheme:



The physical meaning of this scheme is self-evident.

There are four types of 'transitions' which to some extent furnish a simple analogue to the 'transitions' postulated hitherto in the theory of the hydrogen atom, viz. (1) from ellipse to ellipse; (2) from ellipse to hyperbola; (3) from hyperbola to ellipse; (4) from hyperbola to hyperbola.

One can still raise the objection against the formulae (38) and (40), ch. 3, that manifestly in some instances the infinite sums on the right-hand sides do not converge, and hence do not represent a function, since of course in classical theory also, the representation of a function $f(p, q)$ by Fourier integrals is sometimes impossible, as for instance if the respective functions $f$ increase linearly with time at large times (as is in general the case with coordinates). To this objection, one may, however, rejoin that the observable effects of the atom (such as radiation, the force upon another atom, etc.) do not in general belong to this type of function, and thus that the appropriate sums of the same type as the formulae (40), ch. 3, might indeed converge.

# CHAPTER 4. PHYSICAL APPLICATIONS OF THE THEORY

## 1. Laws of conservation of momentum and angular momentum; intensity formulae and selection rules

By way of applying the general theory as established in the aforegoing sections, we now derive the known features concerning 'quantization' of angular momentum and some associated principles.

We shall thereby at the same time become acquainted with some characteristic examples involving *integration* of the quantum-mechanical equations of motion. The previously-discussed perturbation methods can, of course, be applied successfully only when a set of particularly simple examples, which can be selected as unperturbed systems $H_0$, has been integrated in some other way. Now, the quantum-mechanical equations of motion coming from the decomposition of matrix equations into components present the special difficulty that – apart from the instance of the harmonic oscillator – infinitely many unknowns occur in each of the separate equations. A technique frequently employed in overcoming this difficulty in the following and, as it seems, of wide applicability, consists of the following procedure: By analogy with classical theory, one first seeks integrals of the equations of motion, i.e., functions $A(p, q)$ which on the basis of the equations of motion and the commutation rules are constant in function of time and consequently become diagonal matrices in the case of nondegenerate periodic systems. Now if $\varphi(p, q)$ be any function whatsoever, the difference

$$\varphi A - A\varphi = \psi$$

can be evaluated with the help of the commutation rules; if $A$ is a diagonal matrix, a system of equations results, each of which contains only a finite number of unknowns, namely a single component of the matrices $\varphi$ and $\psi$ (and two diagonal terms of $A$) in each.

If in Cartesian coordinates, $H = H'(p) + H''(q)$, which includes the case of relativistic mechanics, then one can see immediately that the components of the angular momentum $\mathfrak{M}$, viz.

$$
\begin{aligned}
M_x &= \sum_{k=1}^{f/3} (p_{ky}q_{kz} - q_{ky}p_{kz}),\\
M_y &= \sum_{k=1}^{f/3} (p_{kz}q_{kx} - q_{kz}p_{kx}),\\
M_z &= \sum_{k=1}^{f/3} (p_{kx}q_{ky} - q_{kx}p_{ky})
\end{aligned}
\tag{1}
$$

become constant under the same general conditions as in classical theory. This is because a sum,

$$\dot{M}_z = \varphi(q) + \psi(p),$$

ensues for the derivative of, say, $M_z$ with respect to time, and since all the $p$ commute with one another, as do all the $q$, the quantities $\varphi$, $\psi$ vanish under the same conditions as in classical theory.

The same remarks are to be applied to the linear momentum

$$\mathfrak{p} = \sum_{k=1}^{f/3} \mathfrak{p}_k; \qquad \text{i.e.,} \qquad p_x = \sum_{k=1}^{f/3} p_{kx}, \ldots, \tag{2}$$

which likewise becomes constant. Thus the centre-of-mass theorem holds just as in classical theory.

We immediately note here a formula which will be used later and which can be derived from the commutation relations (3), ch. 2. We find

$$
\begin{aligned}
M_xM_y - M_yM_x &= \sum_{kl} \{(p_{ky}q_{kz} - q_{ky}p_{kz})(p_{lz}q_{lx} - q_{lz}p_{lx})\\
&\qquad - (p_{kz}q_{kx} - q_{kz}p_{kx})(p_{ly}q_{lz} - q_{ly}p_{lz})\},\\
&= \sum_{kl} \{p_{ky}q_{lx}(q_{kz}p_{lz} - p_{lz}q_{kz})\\
&\qquad + q_{ky}p_{lx}(p_{kz}q_{lz} - q_{lz}p_{kz})\},\\
&= \frac{h}{2\pi \mathrm{i}} \sum_{k} (p_{kx}q_{ky} - q_{kx}p_{ky}),
\end{aligned}
$$

i.e.,

$$M_xM_y - M_yM_x = \varepsilon M_z, \qquad (\text{where } \varepsilon = h/2\pi\mathrm{i}). \tag{3}$$

Incidentally, one can directly see from this formula that the theorem of conservation of angular momentum invariably holds for at most one or alternatively for all three axes, as in classical theory.

In the following we shall assume that on treating the problem with which we are confronted by the methods developed in the preceding chapter we are led to obtain *discrete* energy values (point spectrum). If then $\dot{\mathbf{M}}_z=0$ for a *nondegenerate* system – this will for instance be the case if forces which are symmetrical about the $z$-axis act upon the atom – $\mathbf{M}_z$ has to become a *diagonal matrix*: the separate diagonal terms are to be regarded as the angular moments of the atom about the $z$-axis for the individual *states* of the atom. For the investigation of the motions of the electrons in this case, we first note that the relation

$$\mathsf{q}_{lz}\mathbf{M}_z - \mathbf{M}_z\mathsf{q}_{lz} = 0 \tag{4}$$

follows from (1), ch. 4, and since $\mathbf{M}_z(nm)=\delta_{nm}M_{zn}$, this means that

$$q_{lz}(nm)(M_{zn} - M_{zm}) = 0. \tag{5}$$

One sees that: *For a quantum jump in which there is a change in the angular momentum $M_z$, the 'plane of vibration' of the generated 'spherical wave' lies perpendicular to the z-axis.*

Furthermore, one has

$$\begin{aligned} \mathsf{q}_{lx}\mathbf{M}_z - \mathbf{M}_z\mathsf{q}_{lx} &= -\,\varepsilon\mathsf{q}_{ly}, \\ \mathsf{q}_{ly}\mathbf{M}_z - \mathbf{M}_z\mathsf{q}_{ly} &= \quad\varepsilon\mathsf{q}_{lx}, \end{aligned} \tag{6}$$

i.e.,

$$\begin{aligned} q_{lx}(nm)(M_{zn} - M_{zm}) &= -\,\varepsilon q_{ly}(nm), \\ q_{ly}(nm)(M_{zn} - M_{zm}) &= \quad\varepsilon q_{lx}(nm). \end{aligned} \tag{7}$$

*Thus for jumps in which no change in $M_z$ occurs, the emitted light is linearly polarized parallel to the z-axis.*

Further, from (7), ch. 4, it follows that

$$\{(M_{zn} - M_{zm})^2 - (h^2/4\pi^2)\}q_{l\eta}(nm) = 0; \qquad \eta = x, y. \tag{8}$$

One finally concludes: *For every quantum jump $M_{zn}$ changes by 0, or by $\pm h/2\pi$. The light emitted in the latter case is circularly polarized, as follows from (7), ch. 4.*

In accordance with the above finding concerning the possible changes in $M_z$, the quantity $M_{zn}$ can be represented in the form

$$M_{zn} = \frac{h}{2\pi}\,(n_1 + C), \qquad n_1 = \ldots, -2, -1, 0, 1, 2, \ldots. \tag{9}$$

If there were states whose angular momentum did not fit into this set, no transitions and no interactions whatsoever could occur between these and the states depicted by (9), ch. **4**. Equation (9), ch. **4**, can be taken as a motive for splitting $n$ into two components, one of which is the number $n_1$, introduced in (9), ch. **4**, whereas the other, $n_2$, counts off the various $n$ with the same $n_1$. Our matrices then become four-dimensional, and the results we found for the motions of electrons may be summarized as:

$$q_{lz}(nm) = \delta_{n_1, m_1} q_{lz}(nm); \tag{10}$$

$$\begin{aligned} q_{lx}(nm) &= \delta_{1, |n_1 - m_1|} q_{lx}(nm), \\ q_{ly}(nm) &= \delta_{1, |n_1 - m_1|} q_{ly}(nm); \end{aligned} \tag{10'}$$

$$q_{lx}(n_1, n_2; n_1 \pm 1, m_2) \mp \mathrm{i} q_{ly}(n_1, n_2; n_1 \pm 1, m_2) = 0. \tag{10''}$$

Further, from (4) and (6), ch. **4**, it follows that if we set

$$\mathbf{q}_l^2 = q_l^2 = q_{lx}^2 + q_{ly}^2 + q_{lz}^2,$$

then

$$q_l^2 M_z - M_z q_l^2 = 0. \tag{11}$$

This relation means that $q_l^2$ is a diagonal matrix with respect to the 'quantum number' $n_1$.

The relations (4) to (7), ch. **4** and (10), (11), ch. **4**, also hold if in place of the $q_{lx}, q_{ly}, q_{lz}$ we insert $p_{lx}, p_{ly}, p_{lz}$ or alternatively $M_x, M_y, M_z$. Thus in particular we have:

$$\begin{aligned} &M_x(nm) = \delta_{1, |n_1 - m_1|} M_x(nm); \qquad M_y(nm) = \delta_{1, |n_1 - m_1|} M_y(nm), \\ &M_x(n_1, n_2; n_1 \pm 1, m_2) \pm \mathrm{i} M_y(n_1, n_2; n_1 \pm 1, m_2) = 0. \end{aligned} \tag{12}$$

Further (cf. eq. (1), ch. **4**), $\mathfrak{M}^2 = M^2 = M_x^2 + M_y^2 + M_z^2$ is a diagonal matrix with respect to $n_1$, since

$$M^2 M_z - M_z M^2 = 0. \tag{13}$$

For a system in which all three angular momentum conservation theorems apply, the constant components of $\mathfrak{M}$ certainly cannot collectively be diagonal matrices, since otherwise the above considerations for $M_z$ to be a diagonal matrix could be applied to each of these components, which would lead to discrepancies. Hence such a system is necessarily *degenerate.*

We now set out to consider a system $H = H_0 + \lambda H_1 + \ldots$ of the following type: *All three angular momentum theorems are to apply for* $\lambda = 0$. For

$\lambda \neq 0$ *the system is to be nondegenerate; the constancy of* $\mathsf{M}_z$ *is to remain undisturbed. The energy* $\mathsf{H}_0$ *is to be independent of* $n_1$. The results we shall obtain from this investigation of the case $\lambda \neq 0$ can in part also be carried over to the degenerate system $H_0$, namely insofar as *they are independent firstly of* $\lambda$ *and secondly of the distinguished direction* $z$.

The assumed degeneracy of the system for $\lambda = 0$ is expressed by the fact that $\dot{\mathsf{M}}_x$, $\dot{\mathsf{M}}_y$, $(\mathrm{d}/\mathrm{d}t)(\mathsf{M}^2)$ contain no terms of zeroth order in $\lambda$. Thus

$$\begin{aligned} \nu_0(nm)M_\eta(nm) &= 0, \qquad \eta = x, y; \\ \nu_0(nm)M^2(nm) &= 0. \end{aligned} \tag{14}$$

Since $\mathsf{W}_0$ is independent of the quantum number $n_1$ introduced earlier, whence $\nu_0(n_1, n_2; m_1, n_2) = 0$, whereas $\nu_0(n_1, n_2; m_1, m_2) \neq 0$ is invariably non-zero for $n_2 \neq m_2$, it follows from (14), ch. 4, that

$$\begin{aligned} M_\eta^0(nm) &= \delta_{n_2 m_2} M_\eta^0(nm), \\ M^{0^2}(nm) &= \delta_{n_2 m_2} M^{0^2}(nm). \end{aligned} \tag{15}$$

The square of the total momentum $(\mathsf{M}^0)^2$ is a diagonal matrix in consequence of (13), (15), ch. 4. The double sum representing an element of the matrix $\mathsf{M}_x^0$, $\mathsf{M}_y^0$ reduces to a simple sum

$$\begin{aligned} &\sum_{k_1 k_2} M_x^0(n_1 n_2; k_1 k_2) M_y^0(k_1 k_2; m_1 m_2) \\ &\qquad = \delta_{n_2 m_2} \sum_{k_1} M_x^0(n_1 n_2; k_1 n_2) M_x^0(k_1 n_2; m_1 n_2), \end{aligned} \tag{16}$$

which contains only a finite number of summation terms because of the finite number of possible $n_1$ at fixed $n_2$ (the terms of

$$\mathsf{M}^{0^2} = \mathsf{M}_x^{0^2} + \mathsf{M}_y^{0^2} + \mathsf{M}_z^{0^2} \geqslant \mathsf{M}_z^2$$

do not depend on $n_1$). In (3), ch. 4, applied to $\mathsf{M}_x^0$, $\mathsf{M}_y^0$, $\mathsf{M}_z^0$, we can at any given time sum the equations which belong to a given $n_2$ over $n_1$ and obtain,[1] for fixed $n_2$:

$$\sum_{n_1} M_z(n_1 n_2; m_1 n_2) = \sum_{n_1} (n_1 + C) \frac{h}{2\pi} = 0. \tag{17}$$

On noting additionally that, by (12) and (16), ch. 4, the sum (17), ch. 4 vanishes for *every* single uninterrupted sequence of the $n_1$ it follows that at fixed $n_2$ the possible values of $n_1 + C$ form an unbroken series and lie symmetrically with respect to zero. Hence they must

---

[1] In I we already noted that in the case of a finite diagonal sum $D(ab)$ we always have $D(ab) = D(ba)$.

necessarily constitute either *integer* or *half-integer* numbers, the latter being numbers in the series ..., $-\frac{3}{2}$, $-\frac{1}{2}$, $\frac{1}{2}$, $\frac{3}{2}$,.... If for the moment $M_z$ about the $z$-axis we now introduce the notation usually used in the literature, namely $m(h/2\pi)$ in place of $(n_1+C)(h/2\pi)$, this result accordingly shows that the selection rule $m \rightarrow (m+1, m, m-1)$ applies to $m$ and that $m$ is either 'integer' or 'half-integer'.

Our result demonstrates further that exclusion of individual states, such as was, for example, necessary in the past theory of the hydrogen atom in order to prevent collisions between the electron and the nucleus, has no place in the theory proposed here.

We now attempt to derive the selection principle for the 'total momentum quantum number', as also the intensities for the Zeeman effect, from our theory, proceeding from (5) and (8), ch. 4.

Let us recall the derivation of these selection rules in classical theory: There it is only necessary to introduce a coordinate system whose $z$-axis coincides with the direction of the total angular momentum; in the new coordinates the same results can be derived for $\mathfrak{M}$ as were previously obtained for $M_z$. Let us accordingly set up such a coordinate system $x'$, $y'$, $z'$. The relation

$$z' = x\frac{M_x}{M} + y\frac{M_y}{M} + z\frac{M_z}{M}$$

has to hold anyway in order that the $z'$-axis lie in the direction of the total momentum. (In the following, we shall again drop the index $^0$ for simplicity in all momenta and coordinates: the calculations throughout refer to the limiting case $\lambda=0$). Further, we can so arrange it that the $x'$-axis lies in the $x$, $y$-plane. Everything is thereby fixed, and we have

$$x' = y\frac{M_x}{\sqrt{(M_x^2+M_y^2)}} - x\frac{M_y}{\sqrt{(M_x^2+M_y^2)}}$$
$$y' = \frac{z(M_x^2+M_y^2) - xM_zM_x - yM_zM_y}{M\sqrt{(M_x^2+M_y^2)}}.$$

Now let us try a similar procedure in quantum mechanics. We introduce the three quantities

$$\begin{aligned} \mathsf{Z}_l &= q_{lx}\mathsf{M}_x + q_{ly}\mathsf{M}_y + q_{lz}\mathsf{M}_z, \\ \mathsf{X}_l &= q_{ly}\mathsf{M}_x - \mathsf{M}_y q_{lx}, \\ \mathsf{Y}_l &= \mathsf{M}_x q_{lz}\mathsf{M}_x + \mathsf{M}_y q_{lz}\mathsf{M}_y - q_{lx}\mathsf{M}_z\mathsf{M}_x - \mathsf{M}_y\mathsf{M}_z q_{ly}. \end{aligned} \qquad (18)$$

In order to derive the desired selection rules, we still need some commutation relations, which result from (4) and (6), ch. 4 ($\varepsilon = h/2\pi i$):

$$q_{lx}M^2 - M^2 q_{lx} = 2\varepsilon(q_{lz}M_y - M_z q_{ly}) \tag{19}$$

and the equations for $q_{ly}$, $q_{lz}$ which ensue from this on cyclic permutation. It then follows[1] from (3), (4), (6) and (19), ch. 4, that

$$\begin{aligned} X_l M^2 - M^2 X_l &= -2\varepsilon Y_l, \\ Y_l M^2 - M^2 Y_l &= \varepsilon(X_l M^2 + M^2 X_l), \\ Z_l M^2 - M^2 Z_l &= 0. \end{aligned} \tag{20}$$

These equations are fully analogous to the relations (4) and (6), ch. 4, which determine the selection rules for $M_z$; since we shall later show that the $q_{lx}$, $q_{ly}$, $q_{lz}$ really can be expressed as linear functions of the $X_l$, $Y_l$, $Z_l$, with coefficients which for $\lambda = 0$ are constant with time, we can determine the selection rules for $M$ directly from (20), ch. 4. As $M^2$ is a

---

[1] The first and third formulae in eq. (20), ch. 4, result from a quite simple calculation. The second of eqs. (20), ch. 4, can be derived in the following way: From (18), ch. 4,

$$Y_l = M_x q_{lz} M_x + M_y q_{ly} M_y - q_{lx} M_z M_x - M_y M_z q_{ly},$$

and because of (6), ch. 4,

$$\begin{aligned} Y_l &= q_{lz}(M_x^2 + M_y^2) - \varepsilon q_{ly} M_x + \varepsilon M_y q_{lx} + \varepsilon^2 q_{lz} \\ &\qquad - q_{lx} M_z M_x - M_y M_z q_{ly} \\ &= q_{lz}(M^2 - M_z^2) - \varepsilon X_l + \varepsilon^2 q_{lz} - q_{lx} M_z M_x - M_y M_z q_{ly}. \end{aligned}$$

In the evaluation of $Y_l M^2 - M^2 Y_l$ we now have to note that $M^2$ commutes with $M_x$, $M_y$, $M_z$. Hence for the second part of the formula for $Y_l$ written above, it follows that (cf. (19), ch. 4)

$$\begin{aligned} &(q_{lx} M_z M_x + M_y M_z q_{ly})M^2 - M^2(q_{lx} M_z M_x + M_y M_z q_{ly}) \\ &\qquad = 2\varepsilon(q_{lz} M_y M_z M_x - M_z q_{ly} M_z M_x + M_y M_z q_{lx} M_z - M_y M_z M_x q_{lz}). \end{aligned}$$

On noting that (eq. (19), ch. 4) $q_{lz}M^2 - M^2 q_{lz} = 2\varepsilon X_l$, it follows from the commutation relations that

$$\begin{aligned} q_{lz} M_y M_z M_x - M_y M_z M_x q_{lz} &= \varepsilon(M_y M_z q_{ly} - q_{lx} M_z M_x), \\ M_y M_z q_{lx} M_z - M_z q_{ly} M_z M_x &= -X_l . M_z^2 - \varepsilon(M_z q_{ly} M_y - q_{lx} M_x M_z), \end{aligned}$$

and finally we obtain the desired formula (20), ch. 4:

$$\begin{aligned} Y_l M^2 - M^2 Y_l &= 2\varepsilon X_l(M^2 - M_z^2 + \varepsilon^2) - \varepsilon(X_l M^2 - M^2 X_l) + 2\varepsilon X_l M_z^2 \\ &\qquad - 2\varepsilon^2(q_{lx} M_x M_z - q_{lx} M_z M_x + M_y M_z q_{ly} - M_z M_y q_{ly}) \\ &= 2\varepsilon X_l(M^2 - M_z^2 + \varepsilon^2) - \varepsilon(X_l M^2 - M^2 X_l) + 2\varepsilon X_l M_z^2 - 2\varepsilon^3 X_l \\ &= \varepsilon(X_l M^2 + M^2 X_l). \end{aligned}$$

diagonal matrix, it follows from (20), ch. 4, that

$$\begin{aligned} X_l(nm)(M_m^2 - M_n^2) &= -2\varepsilon Y_l(nm), \\ Y_l(nm)(M_m^2 - M_n^2) &= \varepsilon X_l(nm)(M_m^2 + M_n^2), \\ Z_l(nm)(M_m^2 - M_n^2) &= 0. \end{aligned} \tag{21}$$

The last of the eqs. (21), ch. 4, states that no vibrations take place in $Z$ which could entail a change in $M^2$. It follows from the first two equations that

$$X_l(nm)\left\{(M_m^2 - M_n^2)^2 - \frac{h^2}{2\pi^2}(M_m^2 + M_n^2)\right\} = 0. \tag{22}$$

If we now set $M_m^2 = (h/2\pi)^2(a_m^2 - \frac{1}{4})$, where $a_m$ denotes any function of the quantum numbers, eq. (22), ch. 4 yields

$$X_l(nm)\left((a_n - a_m)^2 - 1\right)\left((a_n + a_m)^2 - 1\right) = 0,$$

or, if $X_l(nm)$ does not vanish,

$$a_n = \pm a_m \pm 1. \tag{23}$$

There is no sacrifice of generality in taking $a_m$ as positive and $\geqq \frac{1}{2}$ throughout. The $a_m$ thus constitute a series of the form $C, 1+C, 2+C, \ldots$ where $C$ denotes a constant which is $\geqq \frac{1}{2}$. Setting $a_m = j + \frac{1}{2}$ yields

$$M^2 = j(j+1)(h/2\pi)^2, \tag{24}$$

and the following selection rule holds for $j$:

$$j \rightarrow \begin{cases} j+1 \\ j \\ j-1 \end{cases}.$$

This result is formally reminiscent of the values of $M^2$ which enter the Landé $g$-formula.

If for $M_z$ we now again introduce the designation $m(h/2\pi)$, we find from (12), ch. 4, and the relations

$$\mathbf{M}^2 = \mathbf{M}_x^2 + \mathbf{M}_y^2 + \mathbf{M}_z^2$$

and

$$(\mathbf{M}_x + \mathrm{i}\mathbf{M}_y)(\mathbf{M}_x - \mathrm{i}\mathbf{M}_y) = \mathbf{M}_x^2 + \mathbf{M}_y^2 - \mathrm{i}\varepsilon\mathbf{M}_z = \mathbf{M}^2 - \mathbf{M}_z^2 - \mathrm{i}\varepsilon\mathbf{M}_z$$

that

$$\begin{aligned} &M_x(j, m-1; j, m) + \mathrm{i}M_y(j, m-1; j, m) \\ &\qquad\qquad = \frac{h}{2\pi}\sqrt{(j(j+1) - m(m-1))}, \\ &M_x(j, m; j, m-1) - \mathrm{i}M_y(j, m; j, m-1) \\ &\qquad\qquad = \frac{h}{2\pi}\sqrt{(j(j+1) - m(m-1))}. \end{aligned} \tag{25}$$

For a given value of $j$, the maximum value $m_{\max}$ of $m$ is characterized by the absence of the jumps $m_{\max} \to m_{\max}+1$, i.e., the right-hand side of (24), ch. 4, for example vanishes for such jumps. This gives

$$j = m_{\max}.$$

*Hence $j$ also can be 'integer' or 'half-integer' only.*

The calculation of the intensity formulae for the Zeeman effect, e.g., the dependence of $q_{lx}, q_{ly}, q_{lz}$ upon $m$, now appears very simple. From (18), ch. 4, we derive the relations

$$\begin{aligned} q_{lz} &= (Z_l M_z + \varepsilon X_l + Y_l)M^{-2}, \\ q_{lx} + \mathrm{i}q_{ly} &= [Z_l - q_{lz}(M_z + \mathrm{i}\varepsilon) + \mathrm{i}X_l](M_x - \mathrm{i}M_y)^{-1}, \\ q_{lx} - \mathrm{i}q_{ly} &= [Z_l - q_{lz}(M_z - \mathrm{i}\varepsilon) - \mathrm{i}X_l](M_x + \mathrm{i}M_y)^{-1}, \end{aligned} \tag{26}$$

by solving for $q_{lx}, q_{ly}, q_{lz}$. These equations also furnish the previously postponed proof that the $q_{lx}, q_{ly}, q_{lz}$ can be represented as linear functions of the $X_l, Y_l, Z_l$ with coefficients which for $\lambda=0$ are constant with time. At the same time, eqs. (26) ch. 4, include the desired intensity formulae. This can be seen by first noting that the $X_l, Y_l, Z_l$ are diagonal matrices with respect to $m$, since

$$\begin{aligned} X_l M_z - M_z X_l &= 0, \\ Y_l M_z - M_z Y_l &= 0, \\ Z_l M_z - M_z Z_l &= 0. \end{aligned} \tag{27}$$

Our problem now resolves itself into two parts, namely discussion of intensities for jumps $j \to j$ and $j \to j-1$ (the jumps $j \to j+1$ then do not provide anything new). We first consider the transitions $j \to j$. For these, equation (20), ch. 4 shows that only terms in $Z_l$ are present. We shall call these terms $Z_l(j, m)$. Then, on setting $M_z = m(h/2\pi)$ and taking note of (24), ch. 4, eqs. (26), ch. 4 yield:

$$q_{lz}(j,\, m) = \frac{2\pi}{h}\, Z_l(j,\, m)\, \frac{m}{j(j+1)},$$

$$\begin{aligned}(q_{lx} + \mathrm{i}q_{ly})(j,\, m-1;\, j,\, m)&\\ = \frac{2\pi}{h}\, Z_l(j,\, m-1) \sqrt{\frac{j(j+1) - m(m-1)}{j(j+1)}}&, \qquad (28)\\ (q_{lx} - \mathrm{i}q_{ly})(j,\, m;\, j,\, m-1)&\\ = \frac{2\pi}{h}\, Z_l(j,\, m) \sqrt{\frac{j(j+1) - m(m-1)}{j(j+1)}}&.\end{aligned}$$

Finally, to establish the dependence of the quantity $Z_l(j, m)$ upon $m$, we might use the relation

$$\mathsf{M}_x q_{ly} - q_{ly}\mathsf{M}_x = \varepsilon q_{lz}; \qquad (29)$$

it demonstrates in our case that $Z_l(j, m)$ does not depend on $m$. For the transitions $j \to j$ we thus obtain:

$$\begin{aligned}q_{lz}(j,\, m) &: (q_{lx}+\mathrm{i}q_{ly})(j,\, m-1;\, j,\, m) : (q_{lx}-\mathrm{i}q_{ly})(j,\, m;\, j,\, m-1)\\ &= m : \sqrt{\{j(j+1) - m(m-1)\}} : \sqrt{\{j(j+1) - m(m-1)\}}.\end{aligned} \qquad (30)$$

We treat the jumps $j \to j-1$ analogously. For these, according to (21), ch. 4, we have $X_l(j, m; j-1, m) = (\varepsilon/j)\, Y_l(j, m; j-1, m)$. If, using (26), ch. 4, we express the intensities in terms of $X_l(j, m; j-1, m)$, we obtain:

$$q_{lz}(j,\, m;\, j-1,\, m) = \mathrm{i}\,\frac{2\pi}{h}\, X_l(j,\, m;\, j-1,\, m)\, \frac{1}{j},$$

$$\begin{aligned}(q_{lx} + \mathrm{i}q_{ly})(j,\, m-1;\, j-1,\, m)&\\ = \mathrm{i}\,\frac{2\pi}{h}\, X_l(j,\, m-1;\, j-1,\, m-1)\, \frac{\sqrt{(j-m)}}{j\sqrt{(j+m-1)}}&, \qquad (31)\\ (q_{lx} - \mathrm{i}q_{ly})(j,\, m;\, j-1,\, m-1)&\\ = -\,\mathrm{i}\,\frac{2\pi}{h}\, X_l(j,\, m;\, j-1,\, m)\, \frac{\sqrt{(j+m-1)}}{j\sqrt{(j-m)}}&.\end{aligned}$$

In conclusion, to establish the dependence of the quantity $X_l(j, m; j-1, m)$ upon $m$, we again use the relation (29), ch. 4, which by way of a simple calculation here yields:

$$X_l(j, m; j-1, m) = A(j, j-1)\sqrt{(j^2 - m^2)}. \tag{32}$$

We thus find that

$$\begin{aligned} &q_{lz}(j, m; j-1, m) : (q_{lx} + \mathrm{i}q_{ly})(j, m-1; j-1, m) \\ &: (q_{lx} - \mathrm{i}q_{ly})(j, m; j-1, m-1) = \sqrt{(j^2 - m^2)} : \sqrt{((j-m)(j-m+1))} \\ &\qquad : -\sqrt{((j+m)(j+m-1))}. \end{aligned} \tag{33}$$

The jumps $j \to j+1$ essentially give the same intensities; we here find that

$$\begin{aligned} &q_{lz}(j, m; j+1, m) : (q_{lx} + \mathrm{i}q_{ly})(j, m; j+1, m+1) \\ &: (q_{lx} - \mathrm{i}q_{ly})(j, m+1; j+1, m) = \sqrt{((j+1)^2 - m^2)} \\ &: \sqrt{((j+m+2)(j+m+1))} : -\sqrt{((j-m+1)(j-m))}. \end{aligned} \tag{34}$$

The formulae (30), (33), (34), ch. 4 agree with the intensity formulae derived from correspondence considerations.[1]

We wish just to draw attention to a simple deduction from (21), ch. 4: *The jumps* $\Delta j = 0$ *occur only in the '*$Z_l$*-direction'*. If we consider the motion of a *single* electron about a nucleus, that is, examine the hydrogen atom, it follows directly from (1), ch. 4, that $\dot{Z}$ vanishes. Hence in this case the jumps $\Delta j = 0$ never take place.

## 2. The Zeeman effect

If one carries the Lorentz force $(e/c)[\mathfrak{v}\mathfrak{H}]$ exerted by a magnetic field $\mathfrak{H}$ upon an electron over into quantum mechanics, it seems obvious at first that the normal Zeeman Effect ensues for atoms, since under exactly the same assumptions as are introduced to derive Larmor's Theorem classically for the nuclear atom – namely, neglect of terms with $\mathfrak{H}^2$ – one can derive this theorem here. There is, nevertheless, a certain difference between classical theory and quantum mechanics insofar as the justification for dropping terms in $\mathfrak{H}^2$ is concerned. The neglect of $\mathfrak{H}^2$ in classical theory is certainly permissible for orbits of small dimensions and certainly *impermissible* for very large orbits or indeed, hyperbolic orbits. In quantum mechanics all these orbits, be they the innermost or the outermost, are so closely connected with one another as a result of the kinematics specific to quantum me-

---

[1] S. Goudsmit and R. de L. Kronig, Naturwiss. **13** (1925) 90; H. Hönl, Zs. f. Phys. **32** (1925) 340.

chanics, that indication of the neglect of the quantity $\mathfrak{H}^2$ is not immediately apparent. The probabilities of transitions to free electrons are indeed considerable, even from the ground state.

For an oscillator, we are thus sure of the normal Zeeman effect; on the other hand for the nuclear atom it does not seem to be entirely excluded that the intimate connection between innermost and outermost orbits leads to findings which differ somewhat from the normal Zeeman effect. However, we must emphasize that a whole set of weighty reasons speak against the possibility of explaining the anomalous Zeeman effects on this basis. Rather, one might perhaps hope that the hypothesis of Uhlenbeck and Goudsmit might later provide a quantitative description of the above-mentioned phenomena.

## 3. Coupled harmonic resonators. Statistics of wave fields

A system of coupled harmonic oscillators given by

$$H = \tfrac{1}{2} \sum_{k=1}^{f} \frac{p_k^2}{m_k} + Q(q), \tag{35}$$

with a quadratic form $Q(q)$ of the coordinates (with numerical coefficients) represents the simplest conceivable system having several degrees of freedom. As was established in ch. 2 § 1, the commutation rules remain invariant on simultaneous orthogonal transformation of coordinates and momenta. Therefore, as in classical theory, the system (35), ch. 4, can be transformed into a system of *uncoupled* oscillators. In particular, the vibrations of a crystal lattice can be analyzed into *eigenvibrations*, just as in classical theory. Each individual eigenvibration is to be treated as a simple linear oscillator according to the manner discussed previously in detail, and the synthesis of the various uncoupled oscillators to a single system is to be undertaken in the way explained in ch. 2 § 1. The same will also apply if we go over to the limiting case of a system with infinitely many degrees of freedom and for instance consider the vibrations of an elastic body idealized to a continuum or finally of an electromagnetic cavity.

In the previous quantum theory also, vibrations of an electromagnetic cavity constituted the subject of many detailed investigations, since on the one hand the problem of the harmonic oscillator represents just about the simplest problem which can be treated with the methods used hitherto, and on the other hand the familiar result that the energy of an eigenvibration should be an integer multiple of $h\nu$ exhibits a formal similarity to the fundamental assumptions of the theory of light quanta, so that one might hence expect to gain insight into the nature of light quanta through the consideration of black-body radiation. To be sure, it is clear from the very outset that attacking the problem of light quanta from the above standpoint cannot by any means account for the most important aspect of this problem, namely the phenomenon of coupling of distant atoms, for this problem does not enter at all into the formulation of our questions regarding the vibrations of a cavity. So strong an association between the eigenvibrations of a cavity and the light quanta postulated formerly can nonetheless be drawn that every statistics of cavity eigenvibrations corresponds to a definite statistics of light quanta, and conversely.

Debye[1] has attempted to arrive at such a form of statistics, starting from a distribution of individual light quanta among the eigenvibrations of the cavity. In this manner he was able to derive Planck's formula. However such a mixture of theoretical wave and light-quantum considerations would seem to us hardly to accord with the real nature of the problem. Rather, we believe it to be consistent to separate the theoretical wave-aspect of the problem completely from the theory of light quanta, that is to say, to treat the wave-statistics of black-body radiation throughout by the more general statistical rules applying e.g., to the quantum theory of atomic systems. The statistics applicable to light quanta is then, as we shall show, Bose statistics.[2] This finding hardly seems unnatural, since this statistics has nothing to do with the hypothesis of independent light-corpuscles, but rather is to be regarded as carried over from the statistics of eigenvibrations – which just shows that the assumption of statistically independent light-corpuscles would not meet the case correctly.

However, in each such treatment of cavity radiation by quantum theory hitherto, one encountered the fundamental difficulty that although it led to Planck's law of radiation, it did not yield the correct mean square deviation of energy in an element of volume. One thus finds that a consistent treatment of the natural vibrations of a me-

---

[1] P. Debye, Ann. d. Phys. **33** (1910) 1427; cf. also P. Ehrenfest, Phys. Zs. **7** (1906) 528.

[2] S. N. Bose, Zs. f. Phys. **26** (1924) 178.

chanical system or an electromagnetic cavity in accordance with past theory leads to most serious contradictions. This caused us to hope that the modified kinematics which forms an inherent feature of the theory proposed here would yield the correct value for the interference fluctuations, thus precluding the above contradictions and opening the possibility of setting up a consistent system of statistics for black-body radiation.

The states of the system of oscillators can be characterized by 'quantum numbers' $n_1, n_2, n_3, \ldots$ of the individual oscillators, so that apart from an additive constant the energies of the individual states are given by

$$E_n = h \sum_k \nu_k n_k. \tag{36}$$

The additive constant, the so-called *zero-point energy* is

$$C = \tfrac{1}{2} h \sum_k \nu_k \tag{36'}$$

(in particular, for the limiting case of infinitely many degrees of freedom, it would be infinitely large). From now on, let us simply call the quantity $E_n$ in (36), ch. 4, the *thermal energy*. In accordance with what was stated in Part I, the same statistical weight is to be attached to each of the states of the system characterized by a certain set of values $n_1, n_2, n_3, \ldots$. The consequences of this can immediately be perceived on the basis of the following remark: If waves are propagated with a phase velocity $v$ in an $s$-dimensional isotropic part of space $V = l^s$ the number of *eigenvibrations* for the frequency range $d\nu$ is equal to the number of *'cells'* for $d\nu$ (in the Bose–Einstein sense), and this in fact holds for arbitrary $s$, hence e.g. also for *vibrating membranes* or *strings*. This follows from the fact that, if we omit consideration of polarization properties, etc., the number of eigenvibrations for the range $d\nu$ is furnished by the solution of the problem of determining the number of ways in which one can choose a set of positive integers $m_1, \ldots m_s$ such that the $\nu$ determined by the relation

$$\frac{2l}{v} \nu = \sqrt{(m_1^2 + \ldots + m_s^2)}$$

falls within the interval $d\nu$. If $K_s(a)$ be the volume of an $s$-dimensional sphere of radius $a$, there are $(V/v^s)K_s(\nu)$ eigenvibrations which have a frequency less than $\nu$. On the other hand, the number of cells for the range $d\nu$ can be determined as follows: The momentum components $p_1, \ldots, p_s$ of the quantum satisfy the equation

$$h\nu/v = \sqrt{(p_1^2 + \ldots + p_s^2)},$$

and the size of the cells is $h^s$ in the $2s$-dimensional phase space. One can see from this that the number of cells belonging to a frequency lower than $\nu$ is also equal to $(V/v^s)K_s(\nu)$.

Hence, as mentioned above, one can effect a one-to-one correspondence of cells to eigenvibrations such that the individual pairs always belong the same $d\nu$. This correspondence can, incidentally, be so carried out that the *directions* of an eigenvibration and those of the light quanta in the respective cell fall within the same infinitesimal angular range. From (36), ch. 4, the quantum number of an *oscillator* is then to be set equal to the *number of quanta in the appropriate cell*. Every system of light-quanta statistics yields an associated statistics of natural vibrations and conversely. It can be seen that the statement made above concerning the weighting of the states of the system of oscillators goes directly over into the basic postulate of Bose–Einstein statistics because of this association. The equally probable complexions are defined through a declaration of the number of quanta sitting in each cell.[1]

In Debye statistics, the number of oscillators involving $r$ quanta is (except for a factor which depends on $\nu$ only) equal to

$$\frac{1}{r}\, e^{-r(h\nu/kT)}, \tag{37}$$

and Planck's law arises from

$$\sum_{r=1}^{\infty} e^{-r(h\nu/kT)} = \frac{1}{e^{h\nu/kT} - 1}.$$

It is unsatisfactory that eq. (37), ch. 4, holds only for $r > 0$ and does not also give the number of oscillators involving no quanta. From the new point of view, we have to replace (37), ch. 4, according to

---

[1] A. Einstein, Sitzungsber. d. Preuss. Akad. d. Wiss. (1925) p. 3. Our considerations naturally cannot yield any fresh viewpoint for the valuation of Einstein's hypothesis that this form of statistics is also applicable to an ideal gas.

Bose,[1] by

$$(1 - \mathrm{e}^{-h\nu/kT})\,\mathrm{e}^{-r(h\nu/kT)}, \tag{38}$$

which (to use the terminology of the theory of light quanta) gives the number of '$r$-fold occupied cells', and Planck's formula results from

$$\sum_{r=0}^{\infty} r(1 - \mathrm{e}^{h\nu/kT})\,\mathrm{e}^{-r(h\nu/kT)} = \frac{1}{\mathrm{e}^{h\nu/kT} - 1}.$$

The light-quanta statistics corresponding to Debye's vibration statistics is represented by the theory developed by Wolfke[2] and Bothe.[3] To be sure, these authors do not speak of $r$-fold occupied cells, but designate (37), ch. 4, as the number of '$r$-quantal light-quanta molecules'.

As is known, the above-mentioned shortcomings of classical wave theory become evident in the study of energy deviations in the radiation field as follows: If there is communication between a volume $V$ and a very large volume such that waves having frequencies which lie within a small range $\nu$ to $\nu+\mathrm{d}\nu$ can pass unhindered from one to the other, whereas for all other waves the volumes remain detached, and if $E$ be the energy of the waves with frequency $\nu$ in $V$, then according to Einstein the mean square deviation $\overline{\Delta^2}=\overline{(E-\bar{E})^2}$ can be calculated by an inversion of the Boltzmann Principle. If $z_\nu\,\mathrm{d}\nu$ be the number of eigenvibrations (cells) in the range $\mathrm{d}\nu$ per unit volume, so that

$$\bar{E} = \frac{z_\nu h\nu}{\mathrm{e}^{h\nu/kT} - 1} \cdot V, \tag{39}$$

then it follows that

$$\overline{\Delta^2} = h\nu\bar{E} + \frac{\bar{E}^2}{z_\nu V}. \tag{40}$$

If, however, one calculates the energy deviations from *interferences* in the wave field, classical theory yields only the second summation

---

[1] This expression naturally has to be assumed for example also in the case of elastic waves in a continuum, which necessitates a certain modification to considerations by Schrödinger (Phys. Zs. **25** (1924) 89) concerning the thermal equilibrium between light- and sound-beams. This modification can easily be carried out in analogy with the probability theorem for the Compton effect on assuming Einstein's gas theory to be valid, as has earlier been pointed out (P. Jordan, Zs. f. Phys. **33** (1925) 649).

[2] M. Wolfke, Phys. Zs. **22** (1921) 375.

[3] W. Bothe, Zs. f. Phys. **20** (1923) 145; **23** (1924) 214.

term in (40), ch. 4, as has explicitly been shown by Lorentz.[1] This discrepancy naturally also exists quite generally for such waves as those in a crystal lattice or in an elastic continuum. According to Ehrenfest,[2] its origins are to be sought in the fact that in the Einstein treatment, *additivity of the entropies* of $V$ and of the large volume was assumed. However, this additivity of entropies applies, according to the classical theory of natural vibrations, only in the region of validity of the Rayleigh–Jeans Law. Precisely the nonexistence of statistical independence of the volume elements in the general case is so unnatural a result of the theory of cavity radiation to date that one is obliged to conclude that this theory breaks down even in the simple problem of the harmonic oscillator.

We now calculate the mean square deviation $\overline{\Delta^2}$ from the interferences using quantum mechanics. To avoid calculational complications which have no bearing upon the nature of the case, we base ourselves on the simplest conceivable case, namely a *vibrating string* fastened at its ends. Incidentally, all essential points of the calculation can immediately be taken over in more general instances. We first cite the classical approach.

Let the length of the string be $l$ and its lateral displacement be $u(x, t)$. On introducing the Fourier coefficients $q_k(t)$ as given by

$$u(x, t) = \sum_{k=1}^{\infty} q_k(t) \sin k \frac{\pi}{l} x, \tag{41}$$

or

$$q_k(t) = \frac{2}{l} \int_0^l u(x, t) \sin k \frac{\pi}{l} x \, \mathrm{d}x \tag{41'}$$

as coordinates, the energy of the string goes over into a sum of squares. Namely, for suitable choice of units,

$$H = \frac{1}{2} \int_0^l \left\{ u^2 + \left( \frac{\partial u}{\partial x} \right)^2 \right\} \mathrm{d}x = \frac{l}{4} \sum_{k=1}^{\infty} \left\{ \dot{q}_k(t)^2 + \left( k \frac{\pi}{l} \right)^2 q_k(t)^2 \right\}. \tag{42}$$

---

[1] H. A. Lorentz, Les Théories Statistiques en Thermodynamique (Leipzig, 1916), p. 59.

[2] P. Ehrenfest, Lecture in the Göttingen seminar on the Structure of Matter, Summer 1925. The contents of this lecture were of great assistance to our present considerations. Meanwhile published in Zs. f. Phys. **34** (1925) 362.

More generally, for the energy $E$ in a segment $(0, a)$ of the string, we obtain

$$E = \frac{1}{2}\int_0^a \sum_{j,k=1}^{\infty} \left\{ \dot{q}_j \dot{q}_k \sin j\frac{\pi}{l}x \sin k\frac{\pi}{l}x + q_j q_k jk \left(\frac{\pi}{l}\right)^2 \cos j\frac{\pi}{l}x \cos k\frac{\pi}{l}x \right\} \mathrm{d}x. \tag{43}$$

If in (43), ch. 4, we take only the terms with $j=k$, we find (under the explicit assumption that all wavelengths which come into consideration are small with respect to $a$) just the value $(a/l)H$. From this one sees: The difference

$$\Delta = E - \bar{E},$$

wherein the bar represents an average over the phases $\varphi_k$ in

$$q_k = a_k \cos(\omega_k t + \varphi_k); \qquad \omega_k = k\frac{\pi}{l}, \tag{44}$$

can be derived from (43), ch. 4, by omitting terms of the sum which have $j=k$. This phase average is identical with the time average. On carrying out the integration, one then finds

$$\Delta = \frac{1}{4}\sum_{\substack{j,k=1\\ j\neq k}}^{\infty} \left\{ \dot{q}_j \dot{q}_k K_{jk} + jk q_j q_k \left(\frac{\pi}{l}\right)^2 K'_{jk} \right\}, \tag{45}$$

with

$$\begin{aligned} K_{jk} &= \frac{\sin (j-k)\dfrac{\pi}{l}a}{(j-k)\dfrac{\pi}{l}} - \frac{\sin (j+k)\dfrac{\pi}{l}a}{(j+k)\dfrac{\pi}{l}} \\ &= \frac{\sin(\omega_j - \omega_k)a}{\omega_j - \omega_k} - \frac{\sin(\omega_j + \omega_k)a}{\omega_j + \omega_k}, \\ K'_{jk} &= \frac{\sin (j-k)\dfrac{\pi}{l}a}{(j-k)\dfrac{\pi}{l}} + \frac{\sin (j+k)\dfrac{\pi}{l}a}{(j+k)\dfrac{\pi}{l}} \\ &= \frac{\sin(\omega_j - \omega_k)a}{\omega_j - \omega_k} + \frac{\sin(\omega_j + \omega_k)a}{\omega_j + \omega_k}. \end{aligned} \tag{45'}$$

In consideration of later quantum-mechanical calculations, we write out the mean square deviation $\overline{\Delta^2}$ explicitly. It is

$$\Delta^2 = (\Delta_1 + \Delta_2)^2 = \Delta_1^2 + \Delta_2^2 + \Delta_1\Delta_2 + \Delta_2\Delta_1, \tag{46}$$

with

$$\Delta_1^2 + \Delta_2^2 = \frac{1}{16} \sum_{\substack{j,k=1 \\ j\neq k}}^{\infty} \sum_{\substack{\iota,\varkappa=1 \\ \iota\neq\varkappa}}^{\infty} \Big\{ \dot q_j \dot q_k \dot q_\iota \dot q_\varkappa K_{jk} K_{\iota\varkappa} + jk\iota\varkappa \left(\frac{\pi}{l}\right)^4 q_j q_k q_\iota q_\varkappa K'_{jk} K'_{\iota\varkappa} \Big\}; \tag{46'}$$

$$\Delta_1\Delta_2 + \Delta_2\Delta_1 = \frac{1}{16} \sum_{\substack{j,k=1 \\ j\neq k}}^{\infty} \sum_{\substack{\iota,\varkappa=1 \\ \iota\neq\varkappa}}^{\infty} \left(\frac{\pi}{l}\right)^2 \{ jk q_j q_k \dot q_\iota \dot q_\varkappa K'_{jk} K_{\iota\varkappa} + \iota\varkappa \dot q_j \dot q_k q_\iota q_\varkappa K_{jk} K'_{\iota\varkappa} \}. \tag{46''}$$

Equation (44), ch. 4, implies $\overline{\Delta_1\Delta_2 + \Delta_2\Delta_1} = 0$ and

$$\overline{\Delta^2} = \overline{\Delta_1^2} + \overline{\Delta_2^2} = \frac{1}{8} \sum_{j,k=1}^{\infty} \left\{ \overline{\dot q_j^2}\,\overline{\dot q_k^2} K_{jk}^2 + j^2k^2 \left(\frac{\pi}{l}\right)^4 \overline{q_j^2}\,\overline{q_k^2} K'^2_{jk} \right\}. \tag{47}$$

If we now let the string's length $l$ become very large, the $\omega_k$ get ever closer together, according to (44), ch. 4, so that the sum (47), ch. 4, goes over into an integral:

$$\overline{\Delta^2} = \overline{\Delta_1^2} + \overline{\Delta_2^2} = \frac{1}{8} \int_0^\infty \int_0^\infty d\omega_j \, d\omega_k \, \frac{l^2}{\pi^2} \left\{ \overline{\dot q_j^2}\,\overline{\dot q_k^2} K_{jk}^2 + j^2k^2 \left(\frac{\pi}{l}\right)^4 \overline{q_j^2}\,\overline{q_k^2} K'^2_{jk} \right\}. \tag{47'}$$

Finally, we also assume the 'volume' $a$ to become very large and employ the relation

$$\lim_{a\to\infty} \frac{1}{a} \int_{-\Omega}^{\Omega'} \frac{\sin^2 \omega\alpha}{\omega^2} f(\omega) \, d\omega = \pi f(0) \qquad \text{for} \qquad \Omega, \Omega' > 0. \tag{48}$$

We then see that only the first sum terms $(\sin (\omega_j - \omega_k)a)/(\omega_j - \omega_k)$ in (45), ch. 4, provide an appreciable contribution, and we find for (47'), ch. 4,

$$\overline{\Delta^2} = \frac{al}{8\pi} \int_0^\infty d\omega \{ (\overline{\dot q_\omega^2})^2 + (\omega^2 \overline{q_\omega^2})^2 \}. \tag{49}$$

On the other hand, by (42), ch. 4, the mean energy in the volume $a$ becomes equal to

$$\bar{E} = \frac{a}{l}\cdot\frac{l}{4}\int_0^\infty \mathrm{d}\omega\ \frac{l}{\pi}\cdot\{\overline{\dot{q}_\omega^2} + \omega^2\overline{q_\omega^2}\} = \frac{al}{4\pi}\int_0^\infty \mathrm{d}\omega\{\overline{\dot{q}_\omega^2} + \omega^2\overline{q_\omega^2}\}. \quad (50)$$

Therein we have

$$\overline{\dot{q}_\omega^2} = \omega^2\overline{q_\omega^2}, \quad (51)$$

a relation which, as we would recall, remains valid in quantum theory too, according to ch. 1. In order to obtain the quantities $\overline{\Delta^2}$, $\bar{E}$ employed in (39), (40), ch. 4, we have merely to extract those parts referring to $\mathrm{d}\nu = \mathrm{d}\omega/2\pi$ from (49), (50), ch. 4, and to divide these by $\mathrm{d}\nu$. With $v = a$ we then obtain

$$\overline{\Delta^2} = \frac{\bar{E}^2}{2v}. \quad (52)$$

We see from (44), ch. 4, that in our case $z_\nu = 2$, since

$$\mathrm{d}\omega_k = 2\pi\,\mathrm{d}\nu_k = \frac{\pi}{l}\,\mathrm{d}k.$$

Thence (52), ch. 4, in fact gives precisely the second term in (40), ch. 4.

On going over to quantum mechanics, we have to regard (41, (41′), (42), (43), ch. 4, as matrix equations for $\mathsf{u}$, $\mathsf{H}$, $\mathsf{q}$, $\mathsf{E}$. The quantity $x$, however, remains a number, since if in place of the continuous string we consider an elastic *series of points*, $x$ would denote the *number* (multiplied by the lattice constant) of any given point.

The matrix $\mathbf{q}_k$ has $2f$ dimensions if $f$ be the number of eigenvibrations, i.e., infinitely many in the case of an elastic string. Each of the components $q_k(nm)$ of $\mathbf{q}_k$ vanishes except for those with

$$\left.\begin{aligned} n_j - m_j &= 0 \qquad \text{for} \qquad j \neq k, \\ n_k - m_k &= \pm\,1. \end{aligned}\right\} \quad (53)$$

The phase average of a matrix is that diagonal matrix which coincides with the diagonal of the respective matrix. From (53), ch. 4, in part similar conclusions can be drawn to those derivable from (44), ch. 4. The considerations which formerly led to (46), (46′), (46″), ch. 4, remain valid in quantum theory. The formulae (47), (47′), ch. 4 with matrices

$\mathbf{q}_k$ also hold for the diagonal matrix $\overline{\boldsymbol{\Delta}_1^2+\boldsymbol{\Delta}_2^2}$ and finally, according to (52), ch. 4, if we denote those parts of $\overline{\boldsymbol{\Delta}^2}$ which belong to a given frequency $\nu$ as $\overline{\Delta^2}$, we find

$$\overline{\Delta_1^2 + \Delta_2^2} = \frac{\overline{E^{*2}}}{2v}. \tag{52'}$$

The quantity $E^*$ in (52′), ch. 4, is, by (49), (50), (51), ch. 4, no longer the mean *thermal* energy, but rather the sum of this and the *zero-point energy*: from the elementary oscillator formulae, we have

$$\overline{E^*} = h\nu \cdot V + \bar{E},$$

$$\overline{\Delta_1^2 + \Delta_2^2} = \tfrac{1}{2}(h\nu)^2\, V + h\nu\bar{E} + \frac{\overline{E^2}}{2V}, \tag{54}$$

since for d$\nu$ the zero-point energy becomes equal to

$$\frac{v}{l}\cdot\frac{h\nu}{2}\cdot lz_\nu\,\mathrm{d}\nu = h\nu\cdot V\,\mathrm{d}\nu.$$

We now still have to consider $\boldsymbol{\Delta}_1\boldsymbol{\Delta}_2+\boldsymbol{\Delta}_2\boldsymbol{\Delta}_1$. In treating this quantity in just the same way as $\overline{\boldsymbol{\Delta}_1^2+\boldsymbol{\Delta}_2^2}$ we obtain, in accordance with (49), ch. 4, the expression:

$$\overline{\boldsymbol{\Delta}_1\boldsymbol{\Delta}_2 + \boldsymbol{\Delta}_2\boldsymbol{\Delta}_1} = \frac{al^2}{8\pi}\int_0^\infty \mathrm{d}\omega\cdot\omega^2\{(\mathbf{q}_\omega\dot{\mathbf{q}}_\omega)^2 + (\dot{\mathbf{q}}_\omega\mathbf{q}_\omega)^2\}.$$

However, since the quantity $\frac{1}{2}l$ is, from (42), ch. 4, to be regarded as the 'mass' of the resonators, the commutation rules give us

$$-\mathbf{q}_j\dot{\mathbf{q}}_j(nn) = \dot{\mathbf{q}}_j\mathbf{q}_j(nn) = \frac{1}{2}\cdot\frac{2}{l}\cdot\frac{h}{2\pi\mathrm{i}} = \frac{h}{2l\pi\mathrm{i}}.$$

Hence the part $\overline{\Delta_1\Delta_2+\Delta_2\Delta_1}$ of $\overline{\boldsymbol{\Delta}_1\boldsymbol{\Delta}_2+\boldsymbol{\Delta}_2\boldsymbol{\Delta}_1}$ which belongs to d$\boldsymbol{\nu}$ is, after division by d$\nu$, equal to

$$\overline{\Delta_1\Delta_2 + \Delta_2\Delta_1} = -\tfrac{1}{2}(h\nu)^2 V,$$

and, with (54), ch. 4, we have in fact

$$\overline{\Delta^2} = h\nu\bar{E} + \frac{\overline{E^2}}{z_\nu V}, \tag{55}$$

which agrees with (40), ch. 4.

If one bears in mind that the question considered here is actually somewhat remote from the problems whose investigation led to the growth of quantum mechanics, the result (55), ch. 4, can be regarded as particularly encouraging for the further development of the theory.

From Ehrenfest's finding mentioned above, one could save oneself calculation of energy deviations involving interference considerations and at the same time acquire the assurance that also for other similar problems no contradictions are possible – if the additivity of the entropies of volume elements could directly be proved in the quantum mechanics of wave fields. Our above findings lead us to expect this additivity to hold generally.

The reasons leading to the appearance in (55), ch. 4, of a term which is not provided by the classical theory are obviously closely connected with the reasons for occurrence of a zero-point energy. The basic difference between the theory proposed here and that used hitherto in both instances lies in the characteristic kinematics and not in a disparity of the mechanical laws. One could indeed perceive one of the most evident examples of the difference between quantum-theoretical kinematics and that existing hitherto on examining the formula (55), ch. 4, which actually involves no mechanical principles whatsoever.

If the quantum mechanics proposed here should prove to be correct even in its essential features, we might quite generally designate the following as constituting the most important advance of this as against the past theory: that in our theory, kinematics and mechanics have again been brought into as close a relationship as that prevailing in classical theory, and that the new fundamental viewpoints, stemming as they do from the basic postulates of quantum theory for the mechanical concepts together with the concepts of space and time, find adequate expression in kinematics just as in mechanics and in the connection between kinematics and mechanics.