OPTICS, MECHANICS, AND WAVE MECHANICS

By

A. LANDÉ

In Tübingen

With 6 Figures

Translated by D. H. Delphenich

From *Handbuch der Physik*, v. 20, L. Grebe, *et al.* (eds.), Springer, Berlin, 1928, pp. 317-452.

TABLE OF CONTENTS

Chapter V – Interaction of Radiation and Matter

Chapter VI – Relativistic Wave Mechanics

Chapter VII – Quantum Algebra and Transformations

INTRODUCTION

1. Overview. – **Newton**'s emission theory leads to the problem of studying not only the rectilinear paths of light corpuscles in a vacuum, but also the curved paths that light will follow in an inhomogeneous anisotropic medium. All possible ray paths shall then be derived from certain optical properties of the medium that generally vary with position and the direction of the ray. According to geometrical optics, only a single quantity namely, the *index of refraction* n, must be known at each location and in any direction at that place in order to establish the ray-curve between any two arbitrary points. (Variations of color and polarization are ignored in that.) Indeed, light seeks the path between two points P and P' along which the *optical length:*

$$S = \int_{P}^{P'} n\,ds$$

assumes an extreme value (**Fermat**'s principle). **Fermat**'s principle does justice to diffraction phenomena and all of the phenomena that are peculiar to the wave theory of light only approximately, and only as long as the index of refraction varies vanishingly little along a wavelength. In that way, e.g., discontinuities, diffraction gratings, etc., are excluded.

Based upon the emission theory, the question arose of whether and how the path of a light ray could be reduced to *forces* that originate in the medium and compel the light corpuscles to follow a curved path. Such a *mechanics of light* should explain the elementary law of interaction between light and matter, and therefore bring with it an interpretation of the optical index of refraction by reducing the law of the paths (i.e., a variational principle) to fundamental force law (i.e., a differential equation for the family of curves). By pursuing that line of reasoning, **Maupertuis** arrived at his *principle of least action* as the mechanical counterpart to **Fermat**'s principle. **Maupertuis**'s investigations were taken up by **Hamilton** and generalized by him and **Jacobi** to a comprehensive theory of arbitrary mechanical systems. The meaning of the classical investigations for the development of mechanics and astronomy is known (1). The optical applications took a backseat to them, especially since they were more interested in the deviations from geometrical optics, namely the study of waves. Nonetheless, geometrical-optical theorems are still significant when one wishes to illustrate mechanical connections, and the analogy between **Fermat**'s and **Maupertuis**'s principles, between optical and mechanical equations of motion, between **Huygens**'s principle and the **Hamilton-Jacobi** partial differential equation allowed progress in mechanics to benefit indirectly from geometrical optics.

However, **Hamilton**'s school of thought, which was so profound for the structure of mechanics remained totally unused in optics, as a result of the separate development of wave theory. New life was breathed into the old analogy by the quantum-theoretic investigations of **de Broglie** and **Schrödinger**, and once again it was optics that exerted its productive influence on mechanics. The study of blackbody radiation in **Planck**'s *quantum theory* had already led to repercussions in

---

(1) See *Handbuch der Physik*, Bd. V, Chap. 3, "Die Hamilton-Jacobische Theorie der Dynamik," by **L. Nordheim** and **E. Fues.**

classical mechanics and electrodynamics that required a revision that was all the more urgent for the further construction of quantum theory, since the **Rutherford-Bohr** model of the atom preserved classical mechanics only qualitatively. Indeed, the adaptation of the astronomical laws to the atomic microcosmos pointed to the superiority of the **Hamilton-Jacobi** methods of generalization and uniformization, but only up to a limit that was defined by the quanta. The latter could not be organically incorporated into mechanics initially, nor could the fruitful hypotheses by which **Bohr** had blazed new trails into the exploration and interpretation of atomic processes dissuade one from the demand of eventually reconstructing mechanics, in whose fundamentals the quanta were already rooted, rather than adding them afterwards as a restricting auxiliary condition. Here, following a trail that was blazed by de Broglie ([1]), **Schrödinger** ([2]) convincingly sought the assistance of the older **Maupertuis-Hamilton** analogy between optics and mechanics: Just as geometrical optics was built upon the basis of the undulatory theory, the classical "geometrical" mechanics was extended to an undulatory mechanics.

That path for getting around the problems of atomic mechanics was all the more surprising since the transition from the discontinuity in the **Newtonian** light corpuscles and the **Einsteinian** light quanta, which are similar to them in many respects, to the continuous wave field of the undulatory theory seemed to point in precisely the opposite direction. However, the spirit of quantum laws was sought in precisely the fact that the continuous state quantities of classical mechanics should yield to the quantum discontinuities. In fact, **Heisenberg** ([3]), **Born** and **Jordan** ([4]), and **Dirac** ([5]) constructed a theory of quantum mechanics that represented a true discontinuum theory, in contrast to the classical theory, in which all differential equations were replaced with difference equations, and all observable data were coupled to each other by purely-algebraic (whole number) ones. However, as was shown later, **Schrödinger**'s opposite path, along which one replaces the motions of a discontinuous system of mass-points with f degrees of freedom with a continuous field-like phenomenon in a space of f dimensions, was a representation of one and the same thing that deviated only formally: For **Schrödinger**, the quantum-mechanical problem was the eigenvalue problem for a boundary-value problem in f-dimensional space in which the generally *discrete* eigenvalues are interpreted as the values of the energy for the quantum states of the system. The latter, along with other quantities that are accessible to observation, which take the form of the components of algebraic matrices in **Heisenberg**'s theory, can also be simultaneously represented as coefficients in a development in the eigenfunctions of **Schrödinger**'s differential equation.

The formal connection between the classical mechanics of mass-points and the undulatory mechanics is characterized by the reinterpretation of the impulse p_κ as a *differential operator* by way of conjugate coordinates q_κ:

([1]) **L. de Broglie**, Thèse, Paris 1924; Ann. Phys. (Leipzig) (10) **3** (1925), pp. 22. German version by **K. Becker**.

([2]) **E. Schrödinger**, *Abhandlungen zur Wellenmechanik*, Leipzig 1927; "Quantisierung als Eigenwertproblem," Ann. Phys. (Leipzig) **79** (1926), pp. 361, 489.

([3]) **W. Heisenberg**, "Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen," Zeit. Phys. **33** (1925), pp. 879.

([4]) **M. Born** and **P. Jordan**, "Zur Quantenmechanik," Zeit. Phys. **34** (1925), pp. 858; *ibid.* **35** (1926), pp. 557.

([5]) **P. A. M. Dirac**, Proc. Roy. Soc. **109** (1925), pp. 642; *ibid.* **110** (1926), pp. 561.

$$p_\kappa \to \frac{h}{2i\pi}\frac{\partial}{\partial q_\kappa}, \qquad -E \to \frac{h}{2i\pi}\frac{\partial}{\partial t},$$

when one appeals to **Planck**'s constant *h*, which will turn **Hamilton**'s energy equation:

$$H(q, t, p) - E = 0$$

Schrödinger's partial differential equation:

$$\left\{H\left(q,t,\frac{h}{2i\pi}\frac{\partial}{\partial t}\right)+\frac{h}{2i\pi}\frac{\partial}{\partial t}, \psi(q,t)\right\} = 0$$

for a function ψ of the coordinates *q* and time *t*. Aside from the computational advantages of working with *operators* instead of differential equations, the significance of the operator calculus is in the fact that the ψ-functions themselves take on no direct physical meaning because they are *not invariant* under the group of canonical transformations. Only certain derived functions that we will come to call "matrix elements" have that invariance property and are distinguished by being the physical quantities that are accessible to experiments. That fact is at the basis for the calculus of *quantum mechanics* that **Heisenberg** founded, and **Born** and **Jordan**, as well as **Dirac**, dressed in its adequate mathematical clothing. Starting from the epistemological viewpoint that physical equations should consist of (at least mainly) observable quantities, **Heisenberg**, **Born**, and **Jordan** arrived at a quantum algebra of invariant matrix elements that was derived from the corresponding between the quantities that one would expect from classical mechanics and the data that were actually observed and predicted by quantum theory. However, that matrix algebra had a formal sort of character that suggested an immediate reformulation as an operator calculus. **Born** and **Wiener** (^{1}) recognized that fact, and **Lanczos** (^{2}) was the first to introduce the *state function* as the object of the operators in order to give some physically-intuitive content to the otherwise abstruse operator formalism. **Schrödinger** arrived at his undulatory equation for a state function ψ whose eigen-solutions possess the discreteness that reality demands of its stationary quantum states from an entirely different way of thinking, which was just the analogy with optics.

Although **Schrödinger**'s theory, with its new state function, has the advantage of being more intuitive than the other formalisms, one cannot forget that is has the disadvantage that lies in the *non-invariance* of that state function under the transformations of mechanics (canonical transformations) and would be avoided by calculating with those invariant matrix elements. That relationship between the undulatory mechanics of **Heisenberg**, **Born**, and **Jordan** can be characterized as being similar to the relationship between the theory of the ether and the theory of relativity. The study of the material ether for light as the carrier of the observed field properties had the advantage of its immediate intuitiveness, but the great disadvantage that any moving

---

(^{1}) **M. Born** and **N. Wiener**, "Eine neue Formulierung der Quantengesetze für periodische und nichtperiodische Vorgänge." Zeit. Phys. **36** (1926), pp. 174.

(^{2}) **C. Lanczos**, "Über eine feldmäßige Darstellung der neuer Quantenmechanik," Zeit. Phys. **35** (1926), pp. 812.

coordinate system will require its own special ether. The theory of relativity, which is built upon its requirement of being invariant under Lorentz transformations, must give up the intuitively-graspable character of the ether in favor of many equally-justified space-time systems, but in that way it will reveal the relationships between observable quantities in a formally-transparent way. Similarly, **Schrödinger**'s ψ-function, as an intuitive picture for the appearance of the quantum-theoretic discrete states, is also an immediately powerful heuristic tool for the intuitively-minded physicist. However, since every new canonically-transformed coordinate system has its own ψ-functions, the quantity ψ can be ascribed just as little physical reality as the ether had in the absolute theory.

However, the theories of **Schrödinger** and **Heisenberg**, which seem to be so different on first glance, lead to the same values for the invariant quantities that can be physically interpreted. One can seek their physical interpretation in two directions, namely, as statements with a continuum-hydrodynamical nature or as statements about *the theory of probability and statistics*. The resolution of that dilemma, which can only come from experiments, will undoubtedly lean in favor of the statistical interpretation. In that way, the theory has its starting point in the creation of a theory of waves that is analogous to the undulatory mechanics, which seems quite remote from it. It can be characterized (1) as the study of the *interference* of certain probability functions in a coordinate space as many dimensions as the number of degrees of freedom gives to the system being treated. However, along with that, the new theory has also settled the dispute in optics between the study of continuous waves and the study of discontinuous light quanta by a statistical interpretation of the wave fields as probability field for the appearance of light corpuscles.

---

(1) **P. Jordan**, "Eine neue Begründung der Quantenmechanik," Zeit. Phys. **40** (1926), pp. 809; *ibid.* **44** (1927), pp. 1; **D. Hilbert**, **J. von Neumann**, and **L. Nordheim**, "Über die Grundlagen der Quantenmechanik," Math. Ann. **98** (1927), pp. 1.

CHAPTER I

THE OPTICAL-MECHANICAL ANALOGY (¹).

2. Fermat's principle. – Geometrical optics can be developed from the principle of the distinguished light-path. Here, we shall consider only the special case in which the medium is indeed optically inhomogeneous, but *isotropic*, such that its optical properties are determined by a scalar function of position $n(x_1, x_2, x_3)$ that is called the *index of refraction* in the following way: If a curve:

$$s = \int_P^{P'} ds$$

goes from the point P to the point P' then one denotes the line integral:

$$S = \int_P^{P'} n\,ds = \int_P^{P'} dS \tag{1}$$

along the curve as its *optical length*. Now, **Fermat**'s law of the distinguished light-path says that among the possible paths between P and P', a light ray will choose the one with minimal (or more precisely, extremal) optical length S :

$$0 = \delta S = \delta \int_P^{P'} n\,ds\,. \tag{2}$$

The extremal optical path between two well-defined points in the medium is a function of the two points $S(P, P') = S(P', P)$; we would like to refer to it as the **Fermat** function. If we know the function S, i.e., the value of $S(P, P')$ for every point-pair in the medium then we can derive further geometrical-optical properties of the medium from them.

(¹) **E. T. Whittaker**, *Analytische Dynamik*, German by **Mittelsten-Scheid**. Berlin 1914; **L. Nordheim**, "Prinzipe der Dynamik," *Handbuch der Physik*, Bd. V, Chap. 2; **L. Nordheim** and **E. Fues**, "Hamilton-Jakobische Theorie der Dynamik," *Handbuch der Physik*, Bd. V, Chap. 3. Elaboration on a Hamburg lecture on optics by **W. Lenz**. Furthermore, **L. Flamm**, "Die Grundlagen der Wellenmechanik," Phys. Zeit. **27** (1926), pp. 600.

3. Light rays. – In order to arrive at explicit equations for light rays in a medium with a **Fermat** function $S(P, P')$ that is known to be assumed, we ask which *direction* of the light ray PP' beyond P' will advance to a neighboring point P'' that possesses an optical distance *ds* (*dS* = *n ds*, resp.) from P', so it will have an optical distance of *S* + *dS* from *P* (Fig. 1). From **Fermat**'s principle, P'' will then first lie on the surface about *P* that is drawn at an optical distance of *S* + *dS*, and secondly on the surface about P' that is drawn at an optical distance of *S* + *dS*, so at its contact point. However, since one sees that the smallness of *ds* means that one can regard the function *n* as *constant* in the vicinity of P', the latter surface will be a *sphere* of optical radius *dS* and geometrical radius *ds* = *dS* / *n*. As seen from P', P'' will the lie on the surface *S* + *dS* in the direction of the perpendicular to that surface through P', i.e., in the direction for which the function $S(P, P')$ will experience its maximum increase when one leaves the point P'. The direction of the ray at P' will then point in the direction of the gradient:

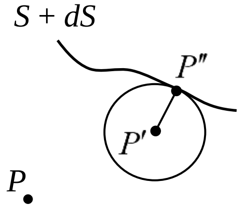

Figure 1.

$$\text{grad}' S = \lambda\,\mathfrak{t}',$$

in which $\mathfrak{t}'$ is a unit vector that is parallel to the light ray, λ is a proportionality factor, and the prime on grad suggests differentiation with respect to the coordinates of P' while *P* is held fixed. Since *dS* = *n ds*, one can determine that $\lambda = n$, and one will get:

$$\text{grad}' S = n\,\mathfrak{t}' \qquad \text{(light ray equation).} \tag{3}$$

Correspondingly, when P', one will get:

$$\text{grad}\, S = -\,n\,\mathfrak{t} \qquad \text{(light ray equation).} \tag{3$'$}$$

at the point *P*, in which the minus sign comes from the fact that *S* will decrease when *P* advances in the direction $\mathfrak{t}$. Each of the two vector equations (3) and (3′) represents three components:

$$\frac{\partial S}{\partial x'} = n\,\mathfrak{t}'_x, \quad \frac{\partial S}{\partial y'} = n\,\mathfrak{t}'_y, \quad \frac{\partial S}{\partial z'} = n\,\mathfrak{t}'_z, \tag{4}$$

$$\frac{\partial S}{\partial x} = -\,n\,\mathfrak{t}'_x, \quad \frac{\partial S}{\partial y} = -\,n\,\mathfrak{t}'_y, \quad \frac{\partial S}{\partial z} = -n\,\mathfrak{t}'_z, \tag{4$'$}$$

in which the left-hand sides will be known functions of the six quantities *x*, *y*, *z*, x', y', z' when the **Fermat** function $S(P, P')$ is assumed to be known. If one then imagines that the six equations have been solved for x', y', z', $\mathfrak{t}'_x$, $\mathfrak{t}'_y$, $\mathfrak{t}'_z$ in the form:

$$x' = x'(x, y, z, \mathfrak{t}_x, \mathfrak{t}_y, \mathfrak{t}_z, S), \quad \ldots, \qquad \mathfrak{t}'_z = \mathfrak{t}'_z\,(x, y, z, \mathfrak{t}_x, \mathfrak{t}_y, \mathfrak{t}_z, S) \tag{5}$$

then one will have a parametric representation (with parameter S) of a light ray that leaves a given point P in a given direction $\mathfrak{t}$. When the Fermat function $S(P, P')$ is assumed to be known, one can then refer to (3) [(3′), resp.] as the implicit *path equations* for the light rays, which correspond to the explicit form (5).

Example: In a homogeneous medium n = const., the **Fermat** function is:

$$S = n\,s = \sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2}\,,$$

so equations (4) will read:

$$n\cdot\frac{x'-x}{s} = n\,\mathfrak{t}'_x\,, \qquad \ldots, \qquad \text{and} \qquad -\,n\cdot\frac{x'-x}{s} = -\,n\,\mathfrak{t}'_x\,, \quad \ldots,$$

and when they are solved one will get:

$$x' = x + s\,\mathfrak{t}'_x = x + \frac{S}{n}\mathfrak{t}'_x\,, \qquad \ldots, \qquad \text{and} \qquad \mathfrak{t}'_x = \mathfrak{t}'_x\,, \qquad \ldots$$

as a special case of (5). The light rays have a constant direction here, so they are straight lines.

The transition from the quantities x_k $\mathfrak{t}_k$ to the x'_k $\mathfrak{t}'_k$ will be given by the transformation (5). It is mediated by the **Fermat** function $S(P, P')$ that appears in (3), (3′) [(4), (4′), resp.] One refers to $S(P, P')$ as the *generator* of the transformation. Due to the contact property above (Fig. 1), one calls the transformation from the old coordinates x_k $\mathfrak{t}_k$ to the new coordinates x'_k $\mathfrak{t}'_k$ that is mediated by S a *contact transformation*. The vectorial form (3), (3′) for the representation of light rays has the advantage that is also valid for arbitrary curvilinear coordinates, as opposed to the **Cartesian** components (4), (4′).

4. Equations of motion of light. – Let the optical length S and the geometric length s along a light ray, as measured from a starting point on the ray, be represented as functions of a *parameter* τ:

$$s = s(\tau)\,, \qquad ds = \dot{s}\,d\tau\,, \qquad S = S(\tau)\,, \qquad dS = \dot{S}\,d\tau\,.$$

When based upon any curvilinear coordinate system, the point of the light ray will possess the following coordinates and their derivatives with respect to τ:

$$x_1(\tau),\ x_2(\tau),\ x_3(\tau)\,; \qquad \dot{x}_1(\tau)\,,\ \dot{x}_2(\tau)\,,\ \dot{x}_3(\tau)\,.$$

$\dot{s}$, as well as $\dot{S}$, will then be a function of the x_k and $\dot{x}_k$:

$$\dot{s}(x,\dot{x}),\ \dot{S}(x,\dot{x}), \qquad \text{which will make} \qquad n\dot{s} = \dot{S}.$$

Fermat's principle (2) can now be written in the form:

$$\delta\int_{\tau_1}^{\tau_2} \dot{S}(x,\dot{x})\,d\tau = 0. \tag{6}$$

The variational problem will be solved by the **Lagrange** equations:

$$0 = \frac{d}{d\tau}\left(\frac{\partial \dot{S}}{\partial \dot{x}_k}\right) - \frac{\partial \dot{S}}{\partial x_k} \qquad \text{for} \qquad k = 1, 2, 3 \tag{7}$$

as the differential equation for $\dot{S}$. Here, one can split off the factor *n*, which depends upon only x_k, but not $\dot{x}_k$, and obtain:

$$\frac{\partial(n\dot{s})}{\partial x_k} = \frac{d}{d\tau}\left(n\frac{\partial \dot{s}}{\partial \dot{x}_k}\right) \qquad \text{for} \qquad k = 1, 2, 3. \tag{8}$$

One can take, e.g., *time t* to be the parameter τ with which one can move a point on the ray (viz., a point of a certain "phase") along the latter. Based upon the velocity law $\dot{s} = c / n$, the **Fermat** principle will become the principle of *shortest light-time:*

$$0 = \delta S = \delta\int n\,ds = \delta\int n\,\dot{s}\,dt = \delta\int c\,dt = c\,\delta t,$$

and (8) will represent the *equations of motion* of the light-point (or light-phase). However, it would actually go beyond the scope of geometrical optics to consider the temporal propagation of a light phase. Rather, here we are concerned with only the geometric form of the light curves that are given by (8) with an arbitrary meaning for τ.

We would now like to try to write (8) vectorially. In order to do that, it suffices to exhibit their vector form when we start from any coordinates, and the most convenient ones are rectilinear. Now, for rectilinear coordinates, we have, in particular:

$$\dot{s} = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2},$$

and

$$\frac{\partial \dot{s}}{\partial x} = 0, \qquad \frac{\partial \dot{s}}{\partial \dot{x}} = \frac{\dot{x}}{\dot{s}} = \frac{dx/d\tau}{ds/d\tau} = \frac{dx}{ds}.$$

When that is substituted in (8), that will give:

$$\dot{s}\,\frac{\partial n}{\partial x} = \frac{d}{d\tau}\left(n\frac{dx}{ds}\right), \ldots$$

and ultimately when it is summarized vectorially:

$$\dot{s}\,\mathrm{grad}\,n = \frac{d}{d\tau}(n\,\mathfrak{t}) \qquad \text{or also} \qquad \frac{d\,(n\,\mathfrak{t})}{ds} = \mathrm{grad}\,n\,, \tag{9}$$

with $\mathfrak{t}$ as the tangential unit vector. That vector equation shows how the direction $\mathfrak{t}$ changes along the ray as a function of n.

In a homogeneous medium, where n = const., (9) will simplify to $d\,\mathfrak{t}\,/\,ds = 0$, i.e., the direction of the ray will not change.

Later on, we will once more find the counterpart to **Newton**'s equations of mechanics (no. **11**).

5. Eikonal surfaces. – If an arbitrary surface F is given in advance then we seek another surface whose points possess the optical distance S from the initial surface F. We would like to denote the initial surface F by S_0 and the desired surface by S. We will get the latter by the following construction (Fig. 2): One lays a surface around each point P of $S = 0$ that has an optical distance of S from P (**Huygens**'s elementary surface and finds the *envelope* of the elementary surfaces. By carrying out the same construction of S_0 with different optical radii S_1, S_2, …, one will get an entire series of *associated light surfaces* (**Huygens**'s envelopes at optical distances of S_1, S_2, … from S_0), that are also referred to as *eikonal surfaces* (¹). Due to **Fermat**'s principle, the contact point P' of the envelope S with the elementary surface of optical radius S about P will then be the point where the optically-shortest ray from P to S pierces S. Instead of starting from S_0 directly, one can also obtain the eikonal surface S_2 with an intermediate stop at S_1 when one performs **Huygens**'s envelope construction with an optical radius of $S_2 - S_1$ when one starts from S_1.

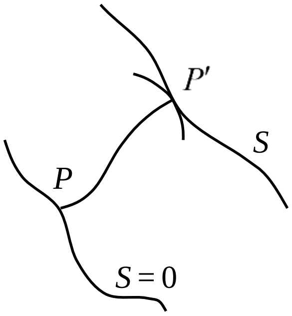

Figure 2.

We would now like to relate a light surface S with the one that is infinitely close to it $S + \delta S$ (Fig. 3). The point P'' at which the ray from P' pierces $S + \delta S$ lies along the direction from P' to $S + \delta S$ that has the shortest optical distance, which is likewise the shortest geometric distance, since index of refraction n is approximately constant there, so the small elementary surface around P' is a *sphere* of geometric radius $\delta s = \delta S\,/\,n$. Therefore, $P'P''$ is *perpendicular* to S and $S + \delta S$. In general, we then have: *The light rays pierce the eikonal surfaces perpendicularly.*

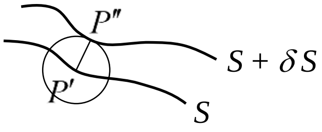

Figure 3.

Along the ray element ds whose direction lies perpendicular to S and $S + \delta S$, i.e., in the direction of the steepest increase (decrease, resp.) of the parameter S, one can write $|\,\delta S\,| : \delta s = |\,\mathrm{grad}\,S\,|$, when $S\,(x, y, z)$ = const. (= S_0, S_1, etc.) represents the equations of

---

(¹) The *eikonal* is an integral that has the value 0 when it is extended between two points of a surface S = const., but the value $S_2 - S_1$ when it is extended between two points of the surfaces S_1 and S_2 .

the eikonal surfaces. Since one has, on the other hand, | δS | : $\delta s = n$, one will get | grad S | = n, which one can also write as:

$$(\text{grad } S)^2 = n^2 \text{ for the eikonal surfaces} \qquad S(x, y, z) = \text{const.} = S\,, \tag{10}$$

or in Cartesian coordinates:

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2\,. \tag{10′}$$

If the family of associated eikonal surfaces is described by the equations $S(x, y, z)$ = const. = S then $S(x, y, z)$ must satisfy the partial differential equation (10). Conversely: If one has a function $S(x, y, z)$ that satisfies the partial differential equation then $S(x, y, z) = S$ will represent a family of associated light surfaces in a medium with an index of refraction n. (10) is the analytical expression for the **Huygens** envelope construction. Later on, we will see that (10) is the optical counterpart of the **Hamilton-Jacobi** partial differential equation for the action function S in mechanics (no. **12**).

6. Light rays and eikonal surfaces. – It shall now be shown that the **Fermat** function $S(P, P')$ in a medium n, as well as the ray-curve between two arbitrary points P and P', can be obtained when one succeeds in finding a solution S to the partial differential equation (grad S)^2 = n^2, and indeed a so-called *complete integral* $S(x, y, z, \alpha_1, \alpha_2, \alpha_3)$ that depends upon constant parameters α_1, α_2, α_3, in addition to the coordinates. If one gives the α fixed values then one will get a family of associated light surfaces of the form:

$$S(x, y, z, \alpha_1, \alpha_2, \alpha_3) = \text{const.} = S \tag{11}$$

with the parameter S. If one gives other fixed values to the α then (11) will represent a different family of associated eikonal surfaces. One can choose the α according to the particular problem.

Now, if $S(x_1, x_2, x_3, \alpha_1, \alpha_2, \alpha_3)$ is a complete integral of (10) then one forms the equations:

$$\frac{\partial S}{\partial \alpha_k} = \beta^k \qquad \text{for} \qquad k = 1, 2, 3 \qquad (\textbf{Jacobi} \text{ system of equations}), \tag{12}$$

in which the β^k are constants that can be once more be adapted to the demands of the particular problem afterwards. Before we prove that, we shall first give an example.

Example: In a homogeneous medium n = const., a complete integral of (10′) is:

$$S = n \cdot \sqrt{(x_1 - \alpha_1)^2 + (x_2 - \alpha_2)^2 + (x_3 - \alpha_3)^2}\ .$$

In place of (12), one will get:

$$\beta^k = -n\,(x_k - \alpha_k) : \sqrt{\cdots} = -\,n^2\,(x_k - \alpha_k) : S$$

here, in particular. It follows that one must choose $\sum(\beta^k)^2 = n^2$, and that:

$$(x_1 - \alpha_1) : (x_2 - \alpha_2) : (x_3 - \alpha_3) = \beta^1 : \beta^2 : \beta^3 .$$

Here, the light rays are lines through the point $x_k = \alpha_k$ with the direction cosines β_k / n.

We can exhibit the general proof that (12) represents light rays when we show the equivalence of (12) with the vector equation grad $S = n\,\mathfrak{t}$ of the light rays, with the use of (grad $S)^2 = n^2$. We next imagine that Cartesian coordinates $\mathfrak{x}_1 = x$, $\mathfrak{x}_2 = y$, $\mathfrak{x}_3 = z$ have been introduced, instead of the arbitrary coordinates x_1, x_2, x_3, and now fix S as the function $S\,(\mathfrak{x}, \alpha)$. Moreover, the parametric representation $\mathfrak{x}_1\,(\tau)$, $\mathfrak{x}_2\,(\tau)$, $\mathfrak{x}_3\,(\tau)$ might be introduced for the curve that is described by (12). We suggest differentiation with respect to τ, i.e., along the curve, by a dot. Since β^i do not dependent upon the $\mathfrak{x}$ (so not upon τ, either), one has:

$$0 = \frac{d\beta^i}{d\tau} = \frac{d}{d\tau}\left(\frac{\partial S}{\partial \alpha_i}\right) = \sum_k \frac{\partial^2 S}{\partial \alpha_i\,\partial \mathfrak{x}_k}\,\dot{\mathfrak{x}}_k .$$

On the other hand, n^2 does not depend upon the α_i, but on the position, such that since $n^2 = (\text{grad}\,S)^2$, one will get:

$$0 = \frac{\partial n^2}{\partial \alpha_i} = \frac{\partial}{\partial \alpha_i}\,(\text{grad}\,S)^2 = \frac{\partial}{\partial \alpha_i}\sum_k \left(\frac{\partial S}{\partial \mathfrak{x}_k}\right)^2 = \sum_k 2\,\frac{\partial S}{\partial \mathfrak{x}_k}\cdot\frac{\partial^2 S}{\partial \mathfrak{x}_k\,\partial \alpha_i} .$$

A comparison of the last two equations will show the proportionalities:

$$\dot{\mathfrak{x}}_k = \lambda\,\frac{\partial S}{\partial \mathfrak{x}_k} \quad \text{and} \quad \dot{s}^2 = \sum_k \dot{\mathfrak{x}}_k^2 = \lambda^2 \sum_k \left(\frac{\partial S}{\partial \mathfrak{x}_k}\right)^2 = \lambda^2\,n^2 ,$$

so $\lambda = \dot{s} / n$ and:

$$\frac{\partial S}{\partial \mathfrak{x}_k} = \frac{\dot{\mathfrak{x}}_k}{\lambda} = \dot{\mathfrak{x}}_k\,\frac{n}{\dot{s}} = n\,\frac{d\mathfrak{x}_k / d\tau}{ds / d\tau} = n\,\mathfrak{t}_k \qquad (\text{for } k = 1, 2, 3),$$

in which $\mathfrak{t}_k$ means the $\mathfrak{x}_k$-component of the unit vector $\mathfrak{t}$ in the direction of the curve. The latter equation can be written vectorially as grad $S = n\,\mathfrak{t}$, so the curve is a light ray. That proves that *the curves* (12) *are light rays.*

One will get the equations of the ray that goes through the two points x_1, x_2, x_3 and x_1', x_2', x_3' when one determines the parameters α_1 to β_3 in (12) with the help of the six equations (k = 1, 2, 3):

$$\frac{\partial S\,(x_1, x_2, x_3, \alpha_1, \alpha_2, \alpha_3)}{\partial \alpha_k} = \beta^k, \qquad \frac{\partial S\,(x_1', x_2', x_3', \alpha_1, \alpha_2, \alpha_3)}{\partial \alpha_k} = \beta^k. \tag{13}$$

The **Fermat** optical distance PP' will then be:

$$S\,(P, P') = S\,(x_1, x_2, x_3, \alpha_1, \alpha_2, \alpha_3) - S\,(x_1', x_2', x_3', \alpha_1, \alpha_2, \alpha_3)\;. \tag{14}$$

7. Phase velocity. – Some further connections between light rays and eikonal surfaces shall now be discussed, but although they belong to the realm of geometrical optics, they are, however, borrowed from the theory of waves, in that the *velocity* of a light ray (the velocity with which a "phase" advances along the light ray, resp.) is drawn into the sphere of consideration.

Let $S\,(x_1, x_2, x_3, \alpha_1, \alpha_2, \alpha_3)$ be a complete integral of:

$$(\text{grad } S)^2 = n^2\,. \tag{15}$$

For fixed α, $S\,(x, \alpha) = S$ will then represent a family of successive eikonal surfaces as the parameter S varies that we call the "α-family." Now, the value of S at time t = 0 shall be referred to as the *phase* of the surface $S\,(x, \alpha) = S$. By contrast, at the time t, the phase of that surface with the value $S - c\,t$:

$$\varphi\,(x, \alpha, t) = S\,(x, \alpha) - c\,t = \text{phase} \tag{16}$$

will define the α-family at the spacetime point (x, t). In that way, c shall possess the value of the speed of light *in vacuo*. A well-defined phase will then wander from surface to surface within the α-family. One can also express that in the form: Every individual surface, along with its phase, wanders throughout the family. Now, since on the one hand, $dS = n\,ds$, and on the other, since (16) ways that $dS = c\,dt$ for constant phase, it will follow that $ds / dt = c / n$ is the velocity with which the phase wanders throughout the α-family and perpendicular to them, or also the phase velocity with which a surface of constant phase moves forward perpendicular to itself:

$$\frac{ds}{dt} = \frac{c}{n} = u = \text{phase velocity.} \tag{16′}$$

Moreover, due to (16) and (16′), one can replace the equation $(\text{grad } S)^2 = n^2$ with the equation for φ:

$$(\text{grad } \varphi)^2 - \frac{1}{u^2}\left(\frac{\partial \varphi}{\partial t}\right)^2 = 0 . \tag{17}$$

From now on, we will consider another family of surfaces, namely, the "$\alpha + \delta\alpha$-family," that differs from the former by only the slightly-varied value of the constant, so the family $S(x, \alpha + \delta\alpha)$ = const. Here as well we define the phase of a surface to be the value of the function:

$$\varphi(x, \alpha + \delta\alpha, t) = S(x, \alpha + \delta\alpha) - c\,t = \text{phase}, \tag{17′}$$

or when developed, since $\partial\varphi / \partial\alpha_k = \partial S / \partial\alpha_k$:

$$\varphi(x, \alpha + \delta\alpha, t) = \varphi(x, \alpha, t) + \sum_k \frac{\partial S}{\partial \alpha_k}\delta\alpha_k = S(x, a) - c\,t + \sum_k \frac{\partial S}{\partial \alpha_k}\delta\alpha_k . \tag{18}$$

At one and the same spacetime point, the α-family will then generate the phase (16), and the $\alpha + \delta\alpha$-family will generate the phase (18). The phase difference D is:

$$\sum_k \frac{\partial S(x, \alpha)}{\partial \alpha_k}\delta\alpha_k = D . \tag{19}$$

At all points that satisfy the latter equation with a given value of D and given quantities $\delta\alpha_k$, the two families α and $\alpha + \delta\alpha$ exhibit the same phase difference D, and in fact at all times, since t does not enter into (19).

Finally, we shall consider yet a third, fourth, etc., family $\alpha + \delta\alpha'$, $\alpha + \delta\alpha''$, etc., which are defined with other well-defined variations $\delta\alpha'$, $\delta\alpha''$, etc., that are, however, all infinitely small. We shall call the entire group of families that are obtained by varying the α-family the α-group. At a spacetime point P, the phase differences at each time t are equal to:

$$D' = \sum_k \frac{\partial S(x^0, \alpha)}{\partial \alpha_k}\delta\alpha_k , \qquad D'_0 = \sum_k \frac{\partial S(x^0, \alpha)}{\partial \alpha_k}\delta\alpha'_k , \qquad \text{etc.} \tag{19″}$$

The condition that the same phase difference should prevail at P as at P^0:

$$D = D_0 , \qquad D' = D'_0 , \qquad D'' = D''_0 , \qquad \dots \tag{20}$$

is fulfilled for:

$$\sum_k \frac{\partial S(x,\alpha)}{\partial \alpha_k}\delta\alpha_k = \sum_k \frac{\partial S(x^0,\alpha)}{\partial \alpha_k}\delta\alpha_k, \qquad \sum_k \frac{\partial S(x,\alpha)}{\partial \alpha_k}\delta\alpha'_k = \sum_k \frac{\partial S(x^0,\alpha)}{\partial \alpha_k}\delta\alpha'_k, \qquad \text{etc.}$$

Since the variations $\delta\alpha,\ \delta\alpha', \ldots$ can possess different infinitely-small values here, the last system of equations is possible only when the equations:

$$\frac{\partial S(x,\alpha)}{\partial \alpha_k} = \frac{\partial S(x^0,\alpha)}{\partial \alpha_k} \qquad \text{for} \qquad k = 1, 2, 3 \tag{20′}$$

are fulfilled individually. All points P that satisfy that system of equations have the property (20), so the same phase difference can prevail there as at P^0. Moreover, one can also write (20′) in the form:

$$\frac{\partial S(x,\alpha)}{\partial \alpha_k} = \beta^k \qquad \text{for} \quad k = 1, 2, 3 \tag{21}$$

and then see from a comparison with (12) that the aforementioned point P is the point of a *light ray* that runs through the α-family perpendicularly and which will naturally include the point P^0 itself, as well. One thus comes to a special conception of light rays: *A light ray, which is perpendicular to the family S* (x, α), *is the locus of points at which the families S* (x, $\alpha + \delta\alpha$), *S* (x, $\alpha + \delta\alpha'$), *etc., possess the same phase differences D,* D', D'', …, *resp., with respect to the family S* (x, α). In that way, the phase was defined by (16), (17′) and shows that it will advance with the velocity $u = c / n$.

8. Group velocity. – The connection between the geometrical-optical results that were discussed up to now and *wave optics* will be exhibited when *one considers periodic functions* ψ *of the phase:*

$$\psi = A \cdot e^{2i\pi\varphi(x,\alpha,t)/\kappa} = A \cdot e^{2i\pi[S(x,\alpha,t)-ct]/\kappa}, \tag{22}$$

in which κ is a constant that should have the same dimension as S. Moreover, one ordinarily refers to the factor of i (so $2\pi S / \kappa$) as the phase, and not S itself. c / k means the temporal frequency ν at a location, and $ds / d(S / \kappa)$ is the length through which one must advance perpendicular to a surface S = const. in order for S / κ to advance by 1:

$$\nu = \frac{c}{\kappa}, \qquad \lambda = \frac{ds\cdot\kappa}{dS}, \qquad \nu\cdot\lambda = c\,\frac{ds}{dS} = \frac{c}{n} = u\,. \tag{22′}$$

We shall now consider the case in which the index of refraction n depends, in addition to position, upon a parameter that we would like to assume is just the frequency $\nu = c / \kappa$: $n = n$ (x, ν). All of the previous relations will likewise include that parameter then. For example, the equation (grad S)^2 = $n^2(x, \nu)$ will possess the wave surfaces S (x, α, ν) = const. as complete integrals whose phases $S = S(x, \alpha, \nu) - c\,t$ will move forward perpendicular to the family of surfaces with a phase velocity $c / n = u(x, \nu)$ that depends upon ν. The periodic functions:

$$\psi(\nu) = A_\nu\, e^{2i\pi\nu[S(x,\alpha,\nu)-ct]/c} \tag{23}$$

will then possess a wavelength that varies with ν. We further consider the light surfaces that belong to the index of refraction $n = n(x, \nu + \delta\nu)$ and their periodic functions:

$$\psi(\nu + \delta\nu) = A_{\nu+\delta\nu}\, e^{2i\pi(\nu+\delta\nu)[S(x,\alpha,\nu+\delta\nu)-ct]/c} . \tag{23′}$$

From the development $S(x, \alpha, \nu + \delta\nu) = S(x, \alpha, \nu) + \dfrac{\partial S}{\partial \nu}\delta\nu$, the *phase difference* Δ between (23) and (23′), i.e., the difference between their exponents, will be equal to:

$$\Delta = \frac{2\pi}{c}\left(\nu\frac{\partial S}{\partial \nu}\delta\nu + S\,\delta\nu - c\,t\,\delta\nu\right) .$$

The phase difference will have the same magnitude for every spacetime point that is described by:

$$\nu\frac{\partial S}{\partial \nu} + S - c\,t = \frac{\partial}{\partial \nu}(\nu\, S) - c\,t = \text{const.} \tag{23″}$$

Now, for a well-defined t, the latter equation will represent a *surface* in space on which the same phase difference will then prevail everywhere between $\psi(\nu)$ and $\psi(\nu + \delta\nu)$. When t increases by δt, the surface of constant phase difference will advance in space, whereby a point P will take on the coordinate increases δx_k, which are restricted by setting the variation of (23″) equal to zero:

$$0 = \sum_k \frac{\partial}{\partial x_k}\left(\frac{\partial}{\partial \nu}(\nu\, S)\right)\cdot\delta x_k - c\cdot\delta t ,$$

or when written vectorially:

$$0 = \left[\text{grad}\left(\frac{\partial \nu\, S}{\partial \nu}\right)\cdot\delta\mathfrak{s}\right] - c\cdot\delta t .$$

In the event that the point P should advance *perpendicular* to the surface (23″), $\delta\mathfrak{s}$ must be taken to be parallel to the grad in the last equation. One will then get the velocity g with which a surface element of the surface of constant phase difference moves forward perpendicular to itself:

$$g = \frac{\delta s}{\delta t} = c : \left| \operatorname{grad} \frac{\partial \nu S}{\partial \nu} \right| ,$$

and g is called the *group velocity:*

$$g = \frac{c}{\left| \operatorname{grad} \frac{\partial \nu S}{\partial \nu} \right|} = \frac{c}{\left| \operatorname{grad} S + \nu \frac{\partial}{\partial \nu} \operatorname{grad} S \right|} . \tag{24}$$

A comparison with the *phase velocity:*

$$u = \frac{c}{n} = \frac{c}{|\operatorname{grad} S|}$$

will give the simple formula:

$$\frac{1}{g} = \frac{\partial \left(\frac{\nu}{u} \right)}{\partial \nu} = \frac{1}{u} \left(1 - \frac{1}{u} \frac{\partial u}{\partial \nu} \right) = \frac{\partial \left(\frac{1}{\lambda} \right)}{\partial \nu} . \tag{24′}$$

An essential path into actual wave optics is that one does not operate with light rays, so one also does not derive the phase relations between the individual spacetime points from their geometrical-optical connection (that is known to lead to false results, such as apparent phase jumps under reflection and diffraction, etc.), but one calculates them from the oscillation function ψ in space and time as a solution to a governing *differential equation*, namely, the wave equation. It is only in the limiting case of small λ that geometrical optics can be applied:

The differential equation of wave optics reads:

$$\Delta \psi - \frac{1}{u^2} \ddot{\psi} = 0 . \tag{25}$$

If one tries to solve it by the Ansatz:

$$\psi = A \cdot e^{2i\pi\varphi/\kappa}$$

then one will get the following differential equation for $\varphi(x, y, z, t)$:

$$\frac{2i\pi}{\kappa} \left[\Delta\varphi - \frac{1}{u^2} \ddot{\varphi} \right] + \left(\frac{2i\pi}{\kappa} \right)^2 \left[(\operatorname{grad} \varphi)^2 - \frac{1}{u^2} \dot{\varphi}^2 \right] = 0 .$$

It is only in the limiting case where the first term will vanish in comparison to the second one (small $\kappa = n\,\lambda$) that the equation of geometrical optics that coincides with (17) will remain:

$$(\operatorname{grad} \varphi)^2 - \frac{1}{u^2} \dot{\varphi}^2 = 0 , \tag{25′}$$

One can now describe the formal connection between (25) and (25′) as follows: One replaces the quantities:

$$\frac{\partial\varphi}{\partial x} \quad \text{with} \quad \frac{\kappa}{2i\pi}\frac{\partial}{\partial x}, \qquad \ldots, \qquad \frac{\partial\varphi}{\partial t} \quad \text{with} \quad \frac{\kappa}{2i\pi}\frac{\partial}{\partial t} \tag{26}$$

in (25′), so, e.g., one replaces $\left(\frac{\partial\varphi}{\partial x}\right)^2$ with $\left(\frac{\kappa}{2i\pi}\frac{\partial}{\partial x}\right)^2 = -\frac{\kappa^2}{4\pi^2}\cdot\frac{\partial^2}{\partial x^2}$, and applies the operator equation that thus arises:

$$\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{u^2}\frac{\partial^2}{\partial t^2} = 0$$

to a function ψ. One will then get the wave equation (25). That formal connection will find its counterpart in the transition from classical to wave mechanics.

9. General metric. – For later use, we will now recall some familiar formulas from the geometric of curvilinear coordinates (¹).

The meaning of the vector operations grad, div, etc., and furthermore perpendicularity, angle, etc., is independent of any coordinate system. If one chooses a certain curvilinear coordinate system x_1, x_2, x_3 as a basis then one can also express *scalar* geometric quantities with the help of components in the system in question. What is characteristic of the system is, above all, the expression for the ordinary length ds of a vector element $d\mathfrak{s}$. $(ds)^2$ is generally a homogeneous quadratic function of the components dx_k :

$$(ds)^2 = \sum_i \sum_k g^{ik}\, dx_i \cdot dx_k\ , \tag{27}$$

with the coefficients g^{ik} that depend upon the position (x_1, x_2, x_3), and which one refers to as the components of the fundamental metric tensor, and in that way, one will have $g^{ik} = g^{ki}$. If one has an *orthogonal* coordinate system then all $g^{ik} = 0$ when $i \neq k$, and only the g^{kk} will be non-zero. In an orthogonal Cartesian coordinate system, the g^{kk} will be constant, and indeed equal to 1, moreover.

Two vectors $d\mathfrak{s}$ and $\delta\mathfrak{s}$ are *perpendicular* to each other when their "scalar product":

$$(d\mathfrak{s},\ \delta\mathfrak{s}) = \sum_i \sum_k g^{ik}\, dx_i \cdot \delta x_k \tag{27.a}$$

vanishes.

(¹) Cf., e.g., **E. Madelung**, *Die mathematischen Hilfsmittel des Physikers*, Berlin, Julius Springer, or **W. Pauli**, *Relativitätstheorie*, Enzykl. d. math. Wiss., Bd. V, 2. Reprinted by B. G. Teubner.

Along with the components dx_k of a vector $d\mathfrak{s}$ that one also calls contravariant components (their index k is attached *below*), one often needs the covariant components dx^k (index above) of the vector $d\mathfrak{s}$, which are defined by:

$$dx^k = \sum_k g^{ik} dx_k \,. \tag{27.b}$$

If one solves the system of equations (27.b) for the dx_k then one will get:

$$dx_k = \sum_i g_{ik}\, dx^k \qquad \text{with the abbreviation } g_{ik} = \frac{\gamma^{ik}}{g^*}, \tag{27.c}$$

in which g^* should mean the determinant of the g^{ik}, and γ^{ik} are the subdeterminants that belong to the g^{ik}. Conversely, one further has:

$$g^{ik} = \frac{\gamma_{ik}}{g_*}, \tag{27.d}$$

in which g_* means the determinant g_{ik}, γ_{ik} is the subdeterminant of g_{ik}.

One then proves the relation:

$$g^* \cdot g_* = 1 \,. \tag{27.e}$$

The *mixed* expressions, in which contravariant and covariant components appear, become especially simple, e.g.:

$$(d\mathfrak{s})^2 = \sum_k dx_k\, dx^k \,, \qquad (d\mathfrak{s}, \delta\mathfrak{s}) = \sum_k dx_k \cdot \delta x^k = \sum_k \delta x_k \cdot dx^k \,.$$

In that way, the g_{ik} drop out completely due to the determinant rule:

$$\sum_i \gamma^{ik} g^{il} = 0 \qquad \text{for} \quad l \neq k, \quad = g^* \quad \text{for } l = k \,.$$

We shall give some further component representations of scalar quantities whose form and magnitude will not change under the transition to other coordinates. The *volume* of a volume element dv is defined by:

$$dv = \sqrt{g^*}\, dx_1\, dx_2\, dx_3 = \sqrt{g_*}\, dx^1\, dx^2\, dx^3 \,. \tag{28}$$

The *gradient vector* of a scalar function of position φ is defined by its components:

$$(\text{grad } \varphi)_k = \frac{\partial \varphi}{\partial x^k} = \sum_i g_{ik} \frac{\partial \varphi}{\partial x_i}, \qquad (\text{grad } \varphi)^k = \frac{\partial \varphi}{\partial x_k} = \sum_i g^{ik} \frac{\partial \varphi}{\partial x^i}, \tag{26.a}$$

from which it will follow that the *gradient-square scalar* is:

$$(\text{grad } \varphi)^2 = \sum_i \sum_k g_{ik} \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_k} = \sum_i \sum_k g^{ik} \frac{\partial \varphi}{\partial x^i} \frac{\partial \varphi}{\partial x^k} . \tag{28.b}$$

The *Laplace operation* Δ = div grad is defined by:

$$\Delta\varphi = \text{div grad } \varphi = \sum_i \sum_k \frac{1}{\sqrt{g^*}} \frac{\partial}{\partial x_i} \left(\sqrt{g^*}\, g_{ik} \frac{\partial \varphi}{\partial x_k} \right) = \sum_i \sum_k \frac{1}{\sqrt{g_*}} \frac{\partial}{\partial x_i} \left(\sqrt{g_*}\, g^{ik} \frac{\partial \varphi}{\partial x^k} \right) . \tag{28.c}$$

All of those formulas can be adapted to n coordinates, instead of three. Moreover, they are true for not only the case in which one deals with *curvilinear* space coordinates $x_1, x_2, \ldots, x_n$, which will take the ordinary (Euclidian) metric to the element of length:

$$ds = \sqrt{\sum_i \sum_k g^{ik}\, dx_i\, dx_k} ,$$

but also for any coordinates in a space in which the element of "length" is defined in a *non-Euclidian* way by the quantity:

$$ds = \sqrt{\sum_i \sum_k g^{ik}\, dx_i \cdot dx_k} .$$

In such a coordinate space, the operations of grad, div, Δ, the concept of perpendicularity, the scalar product, etc., will then have, by definition. the (non-Euclidian) meanings that are adapted from the formulas above. In order to remind ourselves that we are dealing with, say, "perpendicularity" in the non-Euclidian sense of formula (27.a), we will sometimes use "." to emphasize that fact.

10. Maupertuis's and Fermat's principles. – The close analogy between geometrical optics and classical mechanics is based upon the fact that just as optics can be derived from **Fermat**'s principle of the shortest light path, so can mechanics be derived from a variational principle. Ordinarily, one starts from **Hamilton**'s principle in the form:

$$\delta \int_t^{t'} (T - U)\, dt = 0 , \tag{29}$$

in which T is the kinetic energy of the system and U is the potential energy, while t and t' are two time-points, δ is the difference between the actual motion of the system and a neighboring path that has the same initial and final times t and t' and the same initial and final position P and P' that belong to the coordinates.

We would now like to assume that the potential energy U is a function of the N coordinates that describe the position of the system with N degrees of freedom. Furthermore, let the kinetic energy T be a quadratic function of the velocity components $\dot{x}_k$ of the form:

$$T = \tfrac{1}{2} m \sum_i \sum_k g^{ik} \dot{x}_i^2 \, \dot{x}_k^2 , \qquad U = U(x) , \tag{30}$$

in which m is any constant with the dimension of a mass that is included in order for the g^{ik} (which can themselves be functions of the coordinates) to be *dimensionless*. If one understands ds to be the expression:

$$ds = \sqrt{\sum_i \sum_k g^{ik} dx_i \, dx_k} , \tag{30′}$$

i.e., if one introduces a non-Euclidian metric (no. **9**) into the space of coordinates $x_1, \ldots, x_n$ the T can be written in the form:

$$T = \tfrac{1}{2} m \dot{s}^2 . \tag{31}$$

In mechanics, one derives the constancy of the total energy E during the motion from **Hamilton**’s principle under the assumption that $U = U(x_k)$:

$$T + U = E = \text{const.} \tag{32}$$

Due to (32), one will then have $T - U = 2T - E$, and (29) will become:

$$\delta \int_t^{t'} (2T - E) \, dt = 0 ,$$

which will simplify to the condition:

$$\delta \int_t^{t'} 2T \, dt = 0 , \qquad T = E - U , \tag{32′}$$

due to the facts that E = const. and one has fixed t and t'. One can replace $\sqrt{2T}$ with $\sqrt{m}\,\dfrac{ds}{dt}$ in that, such that the last equation will go to:

$$\delta \int_P^{P'} \sqrt{2m(E - U)} \, ds = 0 \qquad (E = \text{const.}) , \tag{33}$$

which is varied between two fixed initial and final configurations P and P' of the system. The version of **Hamilton**’s principle in (32′), (33) is known as the ***Euler-Maupertuis*** *principle of least* *action*, and is valid under the assumption (30), so it is now completely analogous to **Fermat**'s principle of the distinguished light-path, which likewise requires that one vary with fixed initial and final endpoints. Here, we would like to write it in the form:

$$\delta \int_P^{P'} \frac{dS}{c} = \delta \int_P^{P'} \frac{n\,ds}{c} = \delta \int \frac{1}{u}\,ds\,, \tag{34}$$

in which c is the speed of light *in vacuo* and $u = c / n$ is the phase velocity in the medium. Moreover, we would like to write (33) in the form:

$$\delta \int_P^{P'} \frac{\sqrt{2m(E-U)}}{E}\,ds = 0 \tag{35}$$

and indicate the analogy between the last two equations with a $\approx$. The optical function n / c of the coordinates corresponds, mechanically, to the function of position $\sqrt{2m(E-U)} / E$ when the total energy E is given any value:

$$\frac{n(x)}{c} \approx \frac{\sqrt{2m(E-U(x))}}{E}. \tag{36}$$

With the use of the analogy (36), we can now translate all optically-derived relationships into mechanical ones without needing any special proofs. The three-dimensional optical coordinate space of x_k (e.g., rectangular coordinates), in which *lengths*, *perpendicularity*, etc., are understood to have their ordinary (i.e., Euclidian) meanings, shall correspond to an N-dimensional coordinate space of the mechanical point-coordinates x_k in which "length," "perpendicularity," etc., are understood to have the non-Euclidian sense of the element of length (30′). In both cases, one has an isotropic index of refraction n (x_k) with the mechanical meaning (36) and distinguishes the trajectories by way of $\delta \int ds = \delta \int n\,ds = 0$, while the action surfaces S = const. that are "perpendicular" to them are distinguished by $(\text{grad } S)^2 = n^2$.

11. Newton's equations of motion. – We next show that the optical equation (9):

$$\text{grad } n = \frac{d}{dt}(n\,\mathfrak{t})$$

($\mathfrak{t}$ = unit vector tangential to the ray,
$\dot{s}$ = phase velocity = c / n)

is the analogue of **Newton**'s equations of motion for a mechanical system. With the use of the analogy (36), we will, in fact, get from (10) that:

$$\dot{s}\,\mathrm{grad}\sqrt{2m(E-U)} = \frac{d}{dt}\left(\mathfrak{t}\sqrt{2m(E-U)}\right).$$

Due to the fact that $E - U = T = m\dot{s}^2/2$, the left-hand side will be equal to:

$$\dot{s}\,\mathrm{grad}\,(m\dot{s}) = \frac{1}{2m}\mathrm{grad}\,(m\dot{s})^2 = \frac{1}{2m}\mathrm{grad}\,[2m(E-U)] = -\,\mathrm{grad}\,U,$$

and the right-hand side:

$$\frac{d}{dt}(\mathfrak{t}\,m\dot{s}) = \frac{d}{dt}(m\dot{\mathfrak{s}}) = m\ddot{\mathfrak{s}},$$

if $\dot{\mathfrak{s}}$ means the velocity vector $\mathfrak{t}\,s$. One will then have:

$$-\,\mathrm{grad}\,U = m\ddot{\mathfrak{s}}, \tag{37}$$

i.e., the equations of motion in the usual vector form force = mass × acceleration. From (28.a), in component form, that reads, more precisely:

$$-\sum_i g_{ik}\frac{\partial U}{\partial x_i} = m\frac{d^2x_k}{dt^2} \qquad (k = 1, 2, \ldots, N) \tag{37′}$$

in arbitrary coordinates x_k, in which the g^{ik} come from the expression (30) for kinetic energy, and g_{ik} are the associated contravariant quantities. If one expresses U as a function of the covariant x^k, instead of the contravariant x_k, viz., $U(x_k) = \bar{U}(x^k)$, then the latter equation can be written more simply as:

$$-\frac{\partial\bar{U}}{\partial x^k} = m\frac{d^2x_k}{dt^2} \qquad (k = 1, 2, \ldots, n). \tag{37″}$$

The rectilinear light rays in a homogeneous medium n = const. correspond to precisely the trajectories of constant potential energy U.

12. Action function and Hamilton-Jacobi equation. – A comparison of (34) and (35) will lead to the following analogy: The line element of the "optical length" $dS = n\,ds$ will correspond mechanically to the quantity:

$$dS = \frac{c}{E}\sqrt{2m(E-U)}\,ds.$$

The "optical distance" S along a ray between two points in a medium with an index of refraction n will then correspond to the quantity:

$$S(x,x') = \int_P^{P'} dS = \frac{c}{E}\int_P^{P'} \sqrt{2m(E-U)}\,ds = \frac{c}{E}\int_P^{P'} \sqrt{2mT}\,ds, \tag{38}$$

along a trajectory between P and P' that is traversed with constant energy. The integral that already appears in the *principle of least action* (33) will be called the *action*. Here, in order to be able to follow through on the optical-mechanical analogy logically, we would like to refer to the integral $S(x,x')$ in (38) that is multiplied by c / E as the *action*, as we can then say that the **Fermat** light-path function of the point-pair PP' is the analogue of the *action function* of mechanics (38) (S has the dimension of length in both cases).

In optics, it is shown that *light rays* are perpendicular to the eikonal surfaces S = const. Here, the trajectories and the surfaces of constant *action* are correspondingly mutually-"perpendicular." Analogous to the situation in optics, an associated family of action surfaces that is "orthogonal" to a family of trajectories in then determined to be the complete integral of the differential equation (10) $(\text{grad } S)^2 = n^2$, which is:

$$(\text{grad } S)^2 = \frac{c^2}{E^2}\cdot 2m(E-U) = \frac{mc^2}{E}\cdot\frac{2T}{E} \tag{39}$$

here. That equation is identical to the ***Hamilton-Jacobi*** *partial differential equation* (H. J. D) in mechanics for the determination of the *action function* S, which one mostly writes in the form tht is solved for E:

$$E = U + T = U(x_k) + \frac{E^2}{2mc^2}(\text{grad } S)^2 \quad [\textbf{Hamilton-Jacobi}]. \tag{39′}$$

[Observe that S has a meaning in (38) that deviates from the usual one.] One arrives at that in mechanics in the following way:

One can infer the covariant *impulse components* $\xi^k = m\dot{x}^k$ from the expression $T = \frac{1}{2}m\sum_i\sum_k g^{ik}\dot{x}_i\,\dot{x}_k$ for the kinetic energy as a function of the velocity components $\dot{x}_k$, since they are defined by the equations:

$$\frac{\partial T}{\partial \dot{x}_k} = m\sum_i g^{ik}\dot{x}_i = m\dot{x}^k = \xi^k, \tag{40}$$

and one can then write T as a function of the impulse components:

$$T = \tfrac{1}{2}m\sum_i\sum_k g^{ik}\dot{x}_i\,\dot{x}_k = \frac{1}{2m}\sum_i\sum_k g_{ik}\,\xi^i\,\xi^k, \tag{40′}$$

and finally replaces the impulse components with derivatives of the action function S in the form:

$$m\dot{x}^i = \xi^i = \frac{E}{c}\cdot\frac{\partial S}{\partial x_i}, \tag{41}$$

which will put T into the form:

$$T = \frac{E^2}{2mc^2}(\text{grad}\,S)^2,$$

which coincides with (39′).

If a *complete integral* S (x_k, α_k) of the H. J. D (39) has been found then one will get the trajectories [just as one gets the rays from (12)] from the equations:

$$\frac{\partial S\,(x_1\alpha)}{\partial\alpha_k} = \beta^k \qquad (k = 1, 2, \ldots), \tag{42}$$

in which the new constants β^k can adapted to the special initial conditions that were imposed upon the trajectory. The equations:

$$\frac{\partial S}{\partial x_i} = \frac{c}{E}\,m\dot{x}^k, \tag{43}$$

which are compatible with (39) (grad $S)^2 = n^2$, will then also determine the velocity $v = \dot{s}$ of the image point on the trajectory. (42), (43) define the ***Jacobi** system of equations* for determining the motion of the system. (43) includes a “local” description of the velocity components that the system can assume *at a location* x that depends upon the integration constants α. By contrast, (42) will then give a “substantial” description of the path that is followed by *a certain* mechanical system that is characterized by the constants α and β.

The temporal advance of the light phase $\varphi = S - c\,t$ corresponds mechanically to such an advance of the phase $\varphi = S - c\,t$ along the action surfaces. φ has the same spatial differential quotients as S, as well as the temporal one $d\varphi / dt = -c$. In summary, one then has:

$$\left.\begin{array}{ll} \dfrac{\partial\varphi}{\partial x_k} = \dfrac{c}{E}m\dot{x}^k, & \dfrac{\partial\varphi}{\partial t} = -c, \\ \dfrac{\partial\varphi}{\partial\alpha_k} = \beta^k, & \dfrac{\partial\varphi}{\partial c} = -t \end{array}\right\} \tag{43′}$$

for the **Jacobi** system of equations, and along with (42), (43), it describes the spatial, as well as temporal, advance of a point on a path with the phase velocity u.

The construction of the family of surfaces S = const. that satisfy the H. J. D. proceeds precisely as in no. **5**: One erects a “perpendicular” on both sides of each point of the initial surface S (x, α) = S_0 with a “length” of:

$$ds = \frac{dS}{n} = dS\,\frac{E}{c\sqrt{2m(E-U)}},$$

and one will then arrive at the two surfaces $S(x, \alpha) = S_0 \pm dS$.

If one proceeds in succession then one will finally arrive at the finitely-distant surfaces $S(x, \alpha) = S_0 \pm \Delta S$. One had also obtained the same result by constructing **Huygens**'s elementary surfaces with the radii:

$$\Delta S = \int_P dS = \int_P \frac{c}{E}\sqrt{2m(E-U)}\,dS = \text{const.}$$

around each point P of the initial surface (the radii are generally curved while the elementary surfaces are small spheres) and constructing their envelopes.

One can also let the surfaces $S(x, \alpha) = \text{const.}$ emerge from each other in time: In optics, the surface elements of the initial surface $S = S_0$ move perpendicularly to themselves with the phase velocity $u = c / n$ and sweeps out the segment:

$$s = \int u\,dt = \int c/n\,dt = \int dS/n$$

in time t, whereby the surface $S(x, \alpha) = S_0$ would become the surface $S(x, \alpha) = S_0 + c\,t$. Just as in mechanics, one can imagine that the action surface $S(x, \alpha) = S_0 + c\,t$ arises from $S(x, \alpha) = S_0$ by displacing the surface element "perpendicularly" (1) to itself during the time t through the "segment":

$$s = \int \frac{c}{n}dt = \int \frac{E\,dt}{\sqrt{2m(E-U)}} = \int \frac{E\,dt}{\sqrt{2mT}} = \int \frac{dS}{n}$$

with the "velocity":

$$u = \frac{c}{n} = \frac{E}{\sqrt{2m(E-U)}} = \frac{E}{\sqrt{2mT}}\,. \tag{44}$$

If one introduces the notation:

$$\varphi(x, \alpha, t) = S(x, \alpha) - c\,t\,,$$

as in no. **7**, in which φ means a complete integral of $(\text{grad}\,\varphi)^2 = n^2$, just like S, then the equation of the time-shifted action surface will be:

$$\varphi(x, \alpha, t) = S_0\,, \tag{44′}$$

in which $\varphi = S_0$ will once more denote its continually-advancing *phase*. That translates the results on families and groups of light-surfaces in nos. **7** and **8** to families and groups of action surfaces, which we can the state without proof in the following section.

---

(1) The expressions in "…" are meant in the sense of the line element (30′) (see also no. **8**). The "…" can then be dropped for a system of N equal masses m and the use of Cartesian coordinates.

13. Action waves. – The geometrical-optical results of nos. **7** and **8** can be adapted to the language of mechanics with no further analysis, and will then read as follows:

One deals with the action function S as a complete integral $S(x, a)$ of the **Hamilton-Jacobi** partial differential equation (H. J. D.) $(\text{grad } S)^2 = n^2$, in the coordinate space with the metric:

$$ds = \sqrt{\sum_i \sum_k g^{ik} dx_i \, dx_k}\,, \qquad \text{where} \qquad T = \tfrac{1}{2} m \sum_i \sum_k g^{ik} \dot{x}_i \, \dot{x}_k \,. \tag{45}$$

A family of action surfaces $S(x, a) = \text{const.}$, which is characterized by a certain choice of the α_k and different values of the const. (we would like to call it the "α-family"), will then correspond to a family of associated eikonal surfaces in an optical medium with an index of refraction:

$$n = \frac{c}{E}\sqrt{2m(E-U)}\ . \tag{46}$$

We assign a *phase* to each surface of the α-family at time t that is defined by the value of $\varphi = S(x, \alpha) - c\,t$:

$$\varphi(x, \alpha, t) = S(x, \alpha) - c\,t = \text{const.} \tag{47}$$

will then mean *one* action surface with constant phase that advances "perpendicularly" to itself with the "velocity" $u = c / n$, so it will successively pass over the positions of the individual surfaces in the α-family.

In addition to the "α-family," we consider the "$\alpha + \delta\alpha$-family," which is given by $S(x$, "$\alpha + \delta\alpha) = \text{const.}$ with well-defined variations $\delta\alpha_k$, but the same α_k as above. Finally, we consider the *entire group* of families $\alpha + \delta\alpha$, $\alpha + \delta\alpha'$, $\alpha + \delta\alpha''$, … The *phases* at the point P at time t are:

$$\varphi(x, \alpha + \delta\alpha, t)\,, \qquad \varphi(x, \alpha + \delta\alpha', t)\,, \qquad \varphi(x, \alpha + \delta\alpha'', t)\,, \qquad \ldots, \tag{47′}$$

and their *phase differences* compare to the phase $\varphi(x, \alpha, t)$ are (19′):

$$\sum_k \frac{\partial S}{\partial \alpha_k}\delta\alpha_k\,, \quad \sum_k \frac{\partial S}{\partial \alpha_k}\delta\alpha'_k\,, \quad \sum_k \frac{\partial S}{\partial \alpha_k}\delta\alpha''_k\,, \quad \ldots, \tag{48}$$

which are independent of t, since time does not enter into $\partial\varphi / \partial\alpha_k = \partial S / \partial\alpha_k$. The corresponding phase differences at the point P_0 at time t_0 are (19″):

$$\sum_k \left(\frac{\partial S}{\partial \alpha_k}\right)_0 \delta\alpha_k\,, \qquad \sum_k \left(\frac{\partial S}{\partial \alpha_k}\right)_0 \delta\alpha'_k\,, \qquad \sum_k \left(\frac{\partial S}{\partial \alpha_k}\right)_0 \delta\alpha''_k\,, \qquad \ldots \tag{48′}$$

Now, should the same phase differences prevail at P as at P_0, then since $\delta\alpha_k$, $\delta\alpha'_k$, $\delta\alpha''_k$ are arbitrary quantities, one must have:

$$\frac{\partial S}{\partial \alpha_k} = \left(\frac{\partial S}{\partial \alpha_k}\right)_0 \qquad (k = 1, 2, 3, \ldots, N) \tag{49}$$

individually, or when written differently:

$$\frac{\partial S(x, \alpha)}{\partial \alpha_k} = \beta^k \qquad (k = 1, 2, 3, \ldots, N). \tag{49$'$}$$

That system of equations establishes a space curve whose points P all exhibit the *same phase difference* between the families of action surfaces α, $\alpha + \delta\alpha$, $\alpha + \delta\alpha'$, … that simultaneously cross through each other there.

However, (49′) is identical to the **Jacobi** system of equations (42), which established optical light rays as mechanical *trajectories* of the image points $x_1, x_2, \ldots, x_N$ of the mechanical system. *One can then regard the trajectories as curves whose points all possess the same phase differences for the families of action surfaces* α, $\alpha + \delta\alpha$, $\alpha + \delta\alpha'$, $\alpha + \delta\alpha''$, … *that simultaneously cross through each other there.*

Whereas the action surface will be pierced by the "perpendicular" trajectories with the phase "velocity" (1):

$$u = \frac{c}{n} = \frac{E}{\sqrt{2m(E - U(x))}}, \tag{50}$$

the image point of the mechanical system will wander along the trajectory with the "velocity":

$$v = \dot{s} = \sqrt{\frac{2T}{m}} = \sqrt{\frac{2(E - U(x))}{m}}. \tag{50$'$}$$

Just as one concerns oneself with *periodic functions* of the phase φ in optics as the4 light surfaces [cf., (22)], i.e., with the *light waves* along the geometrical-optical rays:

$$\psi = A e^{2i\pi\varphi(x,\alpha,t)/\kappa} = A e^{2i\pi[S(x,\alpha) - ct]/\kappa}, \tag{51}$$

so can one also consider corresponding *action waves* in mechanics. In (51), φ will then mean the phase of the action surface [$\varphi = S(x, \alpha) - c\, t$] as a solution to (grad $S)^2 = n^2$, and κ will be a constant with the dimensions of S. Ordinarily, it is not φ that one refers to as the *phase*, but the factor of i in the exponent (so $2\pi\varphi/\kappa$). As in (22′), one will then have:

---

(1) See footnote on pp. 25.

$$\nu = \frac{c}{\kappa}, \qquad \lambda = \frac{ds}{d(S/\kappa)}, \qquad \nu\lambda = \frac{c}{n} = u \tag{52}$$

for the *frequency*, *wavelength*, and *phase velocity* of the ψ-function. [ds shall be "perpendicular" to the surface $S(x, \alpha)$ = const., so it shall point in the direction of the trajectory.] The *phase velocity* is then different from the velocity of the path $v = \dot{s}$.

We would now like to consider the case in which the "index of refraction" n depends upon not only position, but also a *parameter*, which we would again like to assume is the frequency $\nu = c/\kappa$. Since only E is available in n, from (46), E will include the parameter ν, e.g., the phase velocity u of the action waves:

$$u = \frac{c}{n} = \frac{E(\nu)}{\sqrt{2m[E(\nu) - U(x)]}} = \frac{E(\nu)}{\sqrt{2mT}} \qquad \text{with} \quad E = E(\nu) = E\left(\frac{c}{\kappa}\right) \tag{53}$$

will depend upon ν. Since one considers the total energy to be a function of the parameter ν, all relations that are derived with the use of n or E will include the frequency ν. *Inter alia*, the solution S to $(\text{grad } S)^2 = n^2$ will now possess the form $S(x, \alpha, \nu)$, and (51) will go to the periodic function:

$$\psi(\nu) = A_\nu \, e^{2i\pi\nu[S(x,\alpha,\nu) - ct]/c}, \tag{54}$$

which will then lead to a wavelength (52) that depends upon ν. If one now repeats the argument from equation (23) to (24) in no. **8** word-for-word, so one considers not only the action waves (54), but the waves:

$$\psi(\nu + \delta\nu) = A_\nu \, e^{2i\pi(\nu+\delta\nu)[S(x,\alpha,\nu+\delta\nu) - ct]/c}, \tag{54′}$$

and then seeks the surfaces on which *constant* phase differences will exist between $\psi(\nu)$ and $\psi(\nu + \delta\nu)$ at a time t then one will get the following magnitude for the velocity with which such a surface will advance perpendicular to itself, i.e., the *group velocity g* of the action waves [cf., (24), (24″)]:

$$g = \frac{c}{\left|\text{grad}\,\dfrac{\partial(\nu S)}{\partial \nu}\right|} = \frac{1}{\dfrac{\partial(\nu/u)}{\partial \nu}} = \frac{\partial \nu}{\partial(1/\lambda)}. \tag{55}$$

With the use of u from (53), that will assume the form:

$$g = \sqrt{\frac{2[E(\nu) - U(x)]}{m}} = \sqrt{\frac{2T}{m}} = \dot{s} = v. \tag{56}$$

This result, which was discovered by **L. de Broglie** (1) is stated as: *The group velocity g of the action waves is identical to the path velocity v of the mechanical system.*

That theorem is interesting due to the fact that since it suggests that the mechanical motion of a mass-point (the coordinate point in *N*-dimensional space, resp.) can be regarded as the motion of the group maximum of superimposed action waves.

Now, although it is precisely *that* viewpoint that has not proved to be very fruitful in what followed, it is all the more meaningful to pursue the individual action waves with no consideration given to their group composition.

14. De Broglie's phase waves. – The starting point for wave mechanics is take seriously the plan that was initiated by **de Broglie** and followed through by **Schrödinger** with the optical-mechanical analogy that was represented in nos. **10-13**, i.e., to not merely stop with the fact that many of the paths of moving mass-points in mechanics can be represented by the *picture* of surfaces of constant *action* $\varphi(x, t)$ propagating in time t, but to regard that process of oscillation as being itself the true foundation of mechanical phenomena and to glimpse a new *state quantity* in the periodic function $\psi = \begin{matrix}\sin\\ \cos\end{matrix}\, 2\pi\,\varphi(x, t)$, from whose properties the observed facts are derived, and even better, as it would be possible according to classical mechanics or the original quantum theory. The association of the point-mechanical phenomena with the process of the propagation of an action function, as it was developed in no. **13**, is initially extended, in the spirit of quantum theory, by the fact that the relationship between energy E and the frequency of the periodic function that was left open in it is now established by:

$$E = h\,\nu \qquad (h = \textbf{Planck}\text{ constant}). \tag{58}$$

For a mechanical system whose kinetic energy is a *quadratic* function of the velocities and whose potential is any function of the coordinates:

$$U = U(x)\,, \qquad T = \tfrac{1}{2}\, m \sum_i \sum_k g^{ik}\, \dot{x}_i\, \dot{x}_k\,, \tag{59}$$

the *phase velocity u* of the action wave at a location $P(x)$ will be, according to (50):

$$u = \frac{h\nu}{\sqrt{2m[h\nu - U(x)]}}\,, \tag{60}$$

while the *group velocity g* will be given by (56) as:

---

(1) **L. de Broglie**, *Thèse*, Paris, 1924; Ann. de phys. (10) **3** (1925), pp. 22.

$$g = \sqrt{\frac{2[h\nu - U(x)]}{m}} = 1 : \frac{\partial(\nu / u)}{\partial \nu} . \tag{61}$$

In that way, the path length *ds* in the velocity quotient (path-length : time) is measured with the non-Euclidian metric:

$$ds = \sqrt{\sum_i \sum_k g^{ik} dx_i \, dx_k} .$$

Moreover, from (56), the group velocity is equal to the *point velocity v* at the location *P* when the total energy is given as $E = h\,\nu$.

The wavelength λ of the action waves at the location *P* is then calculated to be:

$$\lambda = \frac{u}{\nu} = \frac{h}{\sqrt{2m(E-U)}} = \frac{h}{\sqrt{2mT}} , \tag{62}$$

in which *T* is the kinetic energy that the system would have at the location *P* point-mechanically if its total energy were $E = T + U$. Now, in general, one has $\sqrt{2mT} = p$ = magnitude of the impulse. One shows this as follows: The impulse components p^k are defined by $p^k = \partial T / \partial \dot{x}_k$, i.e., according to (59), $p^k = m \sum_i g^{ki} \dot{x}_i = m \dot{x}^k$. One will then have:

$$| p |^2 = \sum_i \sum_k g_{ik} \, p^i p^k = m^2 \sum_i \sum_k g_{ik} \, \dot{x}^i \dot{x}^k = m^2 \sum_i \sum_k g^{ki} \dot{x}_i \, \dot{x}_k = 2m \, T .$$

In place of (62), one can then write:

$$\lambda = \frac{h}{p} \qquad \text{(\textbf{de Broglie} wavelength).} \tag{63}$$

This **de Broglie** equation is the spatial counterpart to **Planck**’s temporal equation $\nu = E / h$.

Let us consider, say, the example of **Rutherford**’s model for the hydrogen atom with the charge $\pm \varepsilon$. In it, for a circular orbit of radius *a*, one will have:

$$-E = -\frac{U}{2} = T = \frac{\varepsilon^2}{2a} ,$$

so the **de Broglie** wavelength at each location on the circle will be:

$$\lambda = \frac{h}{\sqrt{2mT}} = \frac{h\sqrt{a}}{\sqrt{m\varepsilon^2}} .$$

The periodic function ψ of the azimuth φ (the path-length $a\,\varphi$ along the orbit, resp.) is:

$$\psi = e^{2i\pi(\nu t - \varphi a/\lambda)} .$$

If one demands that ψ is a *single-valued function* of the azimuth, i.e., ψ goes to itself when φ increases by 2π, then that will be possible only when the factor of $i\,\varphi$ in the exponent is equal to a whole number n :

$$n = \frac{2\pi a}{\lambda} = \frac{2\pi a\sqrt{m\varepsilon^2}}{h\sqrt{a}} ,$$

or when solved from a :

$$a = \frac{n^2 h^2}{4\pi^2 m\varepsilon^2} \qquad \text{and} \qquad E = -\frac{\varepsilon^2}{2a} = -\frac{2\pi^2 m\varepsilon^4}{n^2 h^2} .$$

Those are the known formulas for the radius and energy of the n^{th} quantum orbit, which is derived with the help of the quantum conditions $\left(\int p_\varphi \, d\varphi = h\right)$ in **Bohr**'s theory. Here, one will get the allowable energy values E as a consequence of simply the almost-obvious *single-valuedness requirement* for the function ψ of the azimuth φ, instead of the quantum conditions.

We would also like to find the phase waves for the *relativistic* mass-point m that moves in the absence of external forces (inertial motion of an electron). The kinetic energy will no longer be a quadratic form in the velocity components then. According to **de Broglie**, in that case, one initially imagines that the mass-point is at rest in a coordinate system K_0 with the rest-mass m_0 and the relativistic energy $E_0 = m_0 c^2$ and uses the equation:

$$h\,\nu_0 = E_0 = m_0 c^2 \tag{64}$$

to associate the latter with a frequency ν_0 that defines a *synchronous* oscillation $\psi_0 = e^{2\pi\nu_0 t_0}$ of the entire system K_0 at the time-point t_0 . K_0 moves relative to a system K with a velocity of v in the x-direction. In the latter system, the time value is t, where $t_0 = \left(t - \frac{vx}{c^2}\right) : \sqrt{1 - \frac{v^2}{c^2}}$. Therefore, the oscillation above would be exhibited by:

$$\psi = \exp\left[2\pi\nu_0 \cdot \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}\right]$$

relative to K, so one would have a wave in K that advances along x with a:

$$\text{frequency} \qquad \nu = \frac{\nu_0}{\sqrt{1-\frac{v^2}{c^2}}} \quad \text{and} \quad \text{phase velocity } u = \frac{c^2}{v}, \tag{65}$$

and its wavelength in K would prove to be:

$$\lambda = \frac{u}{\nu} = \frac{c^2}{v} \cdot \frac{\sqrt{1-\frac{v^2}{c^2}}}{\nu_0},$$

or when one uses the facts that:

$$\nu_0 = \frac{E_0}{h} = \frac{m_0 c^2}{h} \qquad \text{and} \qquad p = m_0 v : \frac{\sqrt{1-\frac{v^2}{c^2}}}{\nu_0} = \text{relativistic impulse},$$

one would have:

$$\lambda = \frac{h}{p} \qquad (\textbf{de Broglie} \text{ wavelength}), \tag{66}$$

cf., (63). Due to the facts that $E_0 = h\ \nu_0$, $E = h\ \nu = E_0 : \sqrt{1-\frac{v^2}{c^2}} = h\nu_0 : \sqrt{1-\frac{v^2}{c^2}}$, one will have:

$$\text{the point velocity:} \quad v = c\sqrt{1-\frac{\nu_0^2}{\nu^2}},$$

$$\text{the phase velocity:} \quad u = \frac{c^2}{v} = c : \sqrt{1-\frac{\nu_0^2}{\nu^2}}.$$

From that, one calculates the:

$$\text{the group velocity:} \quad g = 1 : \frac{\partial(\nu/u)}{\partial\nu} = c\sqrt{1-\frac{\nu_0^2}{\nu^2}}.$$

One then finds that the group velocity is identical to the point velocity [cf. (56)].

When one addresses the wave function $\psi(x, t)$, which is continuous in space and time, one will first discover the value that one learns to physically assign to ψ itself [functions of ψ, resp. (in the most general sense, e.g., ψ^2, $\partial\psi / \partial x$, $\int \psi\, dx$, etc.)]. We will go into the association of the ψ-functions with observable physical quantities in the following section. Now, as long as one puts the wave function along a ray s of optical length S into the form:

$$\psi = Ae^{2i\pi\varphi(x,t)/\kappa} = Ae^{2i\pi[E(x)-ct]/\kappa} \tag{67}$$

in *optics*, one will remain within the scope of *geometrical optics*, in which the periodicity of the ray has only a secondary significance. However, that will lead to experimentally-correct results only for wide wave-fronts and minor inhomogeneities in the index of refraction. For narrow ray-bundles and strong inhomogeneity (e.g., diffraction from narrow openings, reflection and refraction at discontinuities in the index of refraction), the function sin $(2\pi\,\varphi / \kappa)$ will not lead to correct phases along the light-path. Rather, one must then allow (viz., the **Fresnel-Fraunhoffer** theory) certain "phase shifts" at the diffracting aperture and with reflection from a thin medium in order to get results that are correct to some extent. **Kirchhoff**'s theory first gave exactly correct value to the oscillation quantity ψ by reverting to a differential equation (wave equation) that was imposed upon the function and solving it for the boundary conditions. Things are similar in mechanics. In that case, if one simply sets ψ equal to a periodic function of the phase (*action*, resp.) along a mechanically-possible point-path then one will get a "geometrical" mechanics with periodicity being imposed secondarily (viz., **de Broglie**'s theory). Now, **Schrödinger**'s accomplishment consists of the fact that he recognized the necessity of looking for **de Broglie**'s oscillation function ψ as a solution to a field equation in space and time, and that he also exhibited the fundamental equation of the undulatory mechanics in the special cases and showed how to interpret them physically. As a counterpart to the optical field equation:

$$\Delta\psi - \frac{1}{u^2}\ddot{\psi} = 0 \qquad \text{with} \qquad u = \frac{c}{n(x)},$$

Schrödinger's fundamental equation for undulatory mechanics read:

$$\Delta\psi - \frac{1}{u^2}\ddot{\psi} = 0 \qquad \text{with} \qquad u = \frac{E}{\sqrt{2m[E - U(x)]}}.$$

Quantum theory is introduced into that by the coupling of the constant value of energy E with the temporal periodicity ν of the oscillation function ψ with the help of the equation $E = h\,\nu$.

CHAPTER II

CORPUSCULAR AND WAVE THEORIES OF LIGHT AND MATTER

15. Statistical theory. – The old controversy between the undulatory and emission theories of light, which seems to have long since concluded with a victory for the wave theory, was revived by the observation of optical processes at the quantum level and found its most fruitful expression in **Einstein**'s quantum theory of light. However, it was clear from the outset in that quest that one would no longer have to deal with an either/or situation in the name of the wave and corpuscular theories, but a union of the two that would correctly account for the typical interference processes, as well as the obvious emission and absorption of individual quantum energies and impulses.

An attempt that was aimed along those lines goes back to **Bohr**, **Kramers**, and **Slater**, namely, the *theory of virtual radiation* ([1]). Starting from the fact that the optical observation registered only quantum-theoretic conversions of material particles (retina, film), one would like to regard the radiation that exists between the emission and absorption of particles as a heuristic mathematical fiction for which one can introduce the detailed picture of an intermediate field for the sake of illustration, but without ascribing any "reality" to the same extent as material particles themselves. The reality of the field in the electromagnetic theory of light is based upon the idea that it functions as the carrier of energy, impulse, stress, and current in such a way that the laws of conservation of energy and impulse will be valid for only the system of matter-plus-field. However, since the conservation laws are initially proved empirically for only macroscopic processes. one can regard them as *statistical laws* that are valid for only distributions of a large ensemble of radiation and particles but will break down for the individual reactions.

For example, when a particle A makes a quantum jump to a different atom B with a loss of energy, according to **Bohr**, **Kramers**, **Slater**, there is a *probability* that it will induce an *inverse* quantum jump with a corresponding gain in energy. That probability shall be calculated from the properties of an imaginary classical wave-field that has its source in an atom A, which is envisioned classically, and which corresponds to the actual quantum atom A in some way. However, that "virtual classical field" *allows one only to calculate the transition probability to the atom B* but cannot *appear as the carrier of the energy and impulse itself*, because that virtual field is assumed to be *continuous*, in the classical way, and capable of producing interference, while the absorption and emission of particles has a quantum character. In that way, the laws of conservation of energy and impulse will break down at the detailed level, so there is a probability that B can absorb or emit A immediately, a while later, or even earlier, without the exchanged energy being found anywhere in the meantime.

That picture unites the field and corpuscular theories, because the classical properties of the virtual radiation makes it possible for interference phenomena to take place, but on the other hand, the quantum processes of emission and absorption are allowed, but generally without the conservation of energy and impulse for the system of matter-plus-radiation. However, the weakness of the theory lies precisely in the latter fact, because the experiments of **Bothe** and

---

([1]) **Born**, **Kramers**, and **Slater**, Zeit. Phys, **24** (1924), pp. 69.

Geiger (1), **Compton** and **Simon** (2), **Joffe** and **Dobronrawoff** (3) subsequently showed that the emission from an individual atom A and the absorption by B are by no means coupled statistically, but in a strictly causal way, and indeed in precisely the same way as when a light-quantum flies away from A and in some way creates the corresponding absorption jump in an atom B anywhere while obeying the laws of energy and impulse.

The theory of **Bohr**, **Kramers**, and **Slater** then has a statistical element, which is obviously necessary for the union of the theories of waves and light-quanta but is not introduced at entirely the right point along the way, because, as the experiments show, the light-quanta *exist* in reality as the direct carriers of energy and impulse. By contrast, the *time-point* at which the light-quantum is emitted from A and the direction of the emission is governed by statistics, like the emission of α-rays from radioactive atoms. We will soon learn about a static theory that takes that fact into account, but first we will confine ourselves to the propagation of light *in vacuo*, and only later treat the static theory of the interaction of radiation and material particles (Chap. V) when the quantum theory of radiation is based upon unperturbed matter.

If one restricts oneself to the propagation of light *in vacuo* then initially there will be no possibility of relating **Einstein**'s light quanta with the state quantities for a **Maxwellian** electromagnetic field, because the light-quanta possess no *polarization* and even fewer six-vector properties than the electromagnetic field $\mathfrak{E}_x$, $\mathfrak{E}_y$, $\mathfrak{E}_z$, $\mathfrak{H}_x$, $\mathfrak{H}_y$, $\mathfrak{H}_z$ at least when one excludes artificial hypotheses on their state. Rather, initially the only possibility that exists is that the light-quanta of given energy and impulse are associated with a *scalar field* ψ that is the solution to the equation $\Delta\psi - \frac{1}{u^2}\psi^2 = 0$ in the pre-**Maxwellian** oscillator theory of light (**Kirchhoff**).

Similarly, we shall later encounter the association of material corpuscles with **Schrödinger**'s mechanical oscillations: Material particles also have energy and moment, but no polarization, so they can only be reinterpreted as scalar quantities. In general, according to **Uhlenbeck** and **Goudsmit**, electrons exhibit a rotational impulse whose two allowed direction are very essential for atomic phenomena to take place, and in a certain sense represent the two opposite *polarization states* of matter (4). The scalar theory of light and matter that is considered in the present article must be correspondingly extended to a vector field theory, which is a problem that has already been taken up (Chap. VI).

We shall then stay with the scalar theory of light and consider the association of light-quanta and scalar light oscillations (as a preparation for the association of the corpuscular and wave theories of matter).

16. Association of light quanta and waves. – The field theory of light *in vacuo* determines the field scalar ψ from the equation:

---

(1) **W. Bothe** and **H. Geiger**, Zeit. Phys. **32** (1925), pp. 639.

(2) **Compton** and **Simon**, Phys. Rev. **26** (1925), pp. 289.

(3) **Joffe** and **Dobronrawoff**, Zeit. Phys. **34** (1925), pp. 889.

(4) For the correspondence between polarized light quanta and the magnetic electrons, see especially, **P. Jordan**, Zeit. Phys. **44** (1927), pp. 292. **C. G. Darwin**, Proc. Roy. Soc. **116** (1927), pp. 227.

$$\Delta\psi - \frac{\ddot{\psi}}{c^2} = 0 ,$$

or in rectangular coordinates:

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0 . \tag{1}$$

By contrast, for free-flying light corpuscles, the relativistic energy-impulse connection $p = E / c$, or in rectangular components:

$$- p_x^2 - p_y^2 - p_z^2 + \frac{E^2}{c^2} = 0 , \qquad p = \frac{E}{c} . \tag{2}$$

Those two equations can be formally converted into each other in the following way:

We recall the transition from geometrical to wave optics that was discussed at the conclusion to no. **8**. At that time, the wave equation $\Delta\psi - \ddot{\psi} / u^2 = 0$ was obtained from the geometrical-optical equation (grad $\varphi)^2 - \ddot{\varphi}^2 / u^2 = 0$ by replacing the quantities:

$$\frac{\partial\varphi}{\partial x} \quad \text{with} \quad \frac{\kappa}{2i\pi}\frac{\partial}{\partial x}, \ldots, \frac{\kappa}{2i\pi}\frac{\partial}{\partial t} \tag{3}$$

formally in the latter, and applying the operator thus-obtained:

$$\left(\frac{\kappa}{2i\pi}\right)^2 \left(\frac{\partial^2}{\partial x^2} + \cdots - \frac{1}{u^2}\frac{\partial^2}{\partial t^2}\right)$$

on a function ψ. (κ stems from the Ansatz $\psi = e^{2i\pi\varphi/\kappa}$, which satisfies the wave equation only approximately, and has the meaning of $\kappa = \nu / c$.) Here, we are dealing with special picture for geometrical optics, namely, with light-rays are *mechanical* paths of light corpuscles. However, in mechanics, one gets energy and impulse from the function $\varphi = S - c\,t$ by way of the derivatives [I, (41)]:

$$\frac{E}{c}\frac{\partial\varphi}{\partial x} = p_x , \quad \ldots, \qquad \frac{E}{c}\frac{\partial\varphi}{\partial x} = \frac{E}{c}(-c) = -E ,$$

in which φ does not mean the usual action function of mechanics, c / E times it. The replacement above (3) will then be the following one:

$$p_x \quad \text{with} \quad \frac{\kappa}{2i\pi}\frac{c}{E}\frac{\partial}{\partial x}, \ldots, \qquad (-E) \quad \text{with} \quad \frac{\kappa}{2i\pi}\frac{c}{E}\frac{\partial}{\partial t},$$

or with the use of $\kappa = c / \nu$, $h = E / \nu$:

$$p_x \quad \text{with} \qquad \frac{\kappa}{2i\pi}\frac{c}{E}\frac{\partial}{\partial x}, \ldots, \qquad (-E) \quad \text{with} \quad \frac{\kappa}{2i\pi}\frac{c}{E}\frac{\partial}{\partial t}. \tag{3′}$$

With the help of that reinterpretation, the equation (2) for the impulse and energy of the light-quantum will, in fact, become:

$$\left(\frac{h}{2i\pi}\right)^2\left(-\frac{\partial^2}{\partial x^2}-\cdots+\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)=0\,,$$

which one then applies to a function ψ in order to obtain the wave equation (1) in the form:

$$\frac{h^2}{4\pi^2}\left\{\frac{\partial^2}{\partial x^2}+\cdots-\frac{1}{c^2}\frac{\partial^2}{\partial t^2},\psi\right\}=0\cdot\psi=\{0,\ \psi\}\,. \tag{4}$$

One can also use the notation:

$$\left\{-p_x^2-p_y^2-p_z^2+\frac{E^2}{c^2},\psi\right\}=\{0,\ \psi\} \tag{4′}$$

instead, in which p_x, …, E shall not mean physical quantities now, but the *operators* that were introduced in (3′). The form (4′) for the wave equation then proves its applicability to the corpuscular equation (2) in a particularly instructive way. *The transition from the point theory to the field theory, so from* (2) *to* (4), *is formally characterized by the transition* (3′) *from impulses to operators that are applied to a field function.* The reinterpretation (3) will be inverted in all later cases, so for the transition from point mechanics to wave mechanics.

It is noteworthy that the quantity h that was into the wave equation (4) in the reinterpretation (3) can be dropped on dimensional grounds (which is something that *does not occur* in the corresponding transition to *mechanics*). That is indicated by the fact that for *optics in vacuo,* the classical theory is valid to a greater degree than perhaps in mechanics, but one should not be misled to the conclusion that no physical difference whatsoever exists between the classic and quantum theories of radiation *in vacuo*. That is because **Planck**’s radiation formula already deviates from the classical theory in an essentially *statistical* way.

We shall now consider an Ansatz for a solution to the wave equation (4) that is *periodic* in spacetime, namely:

$$\psi = A\cdot e^{2i\pi(t\,\mathfrak{s}/\lambda-\nu t)} = A\cdot e^{2i\pi\left(\frac{x}{\lambda}\mathfrak{s}_x+\frac{y}{\lambda}\mathfrak{s}_y+\frac{z}{\lambda}\mathfrak{s}_z-\nu t\right)}. \tag{5}$$

The unit vector $\mathfrak{s}$ then gives the wave normal to the monochromatic plane waves. Introducing the Ansatz (5) into the wave equation (4) will the following relation between the constants λ and ν:

$$-\frac{h^2}{\lambda^2}\mathfrak{s}_x^2-\frac{h^2}{\lambda^2}\mathfrak{s}_y^2-\frac{h^2}{\lambda^2}\mathfrak{s}_z^2+\frac{h^2\nu^2}{c^2}=0\,.$$

A comparison with (2) will show that the energy and impulse of a flying light-quantum are associated with the wave quantities λ and ν by the equations:

$$p_x = \frac{h}{\lambda}\mathfrak{s}_x, \ldots, \; E = h\,\nu. \tag{6}$$

The basic reinterpretation (3′) has led us to the **Planck** and **de Broglie** association then. *A monochromatic plane wave corresponds to a current of light-quanta that possesses an energy of* $E = h\,\nu$ *and impulse components of* $p_x = \mathfrak{s}_x \cdot \frac{h}{\lambda}$*, etc.*

Conversely, substituting (6) in (5) will give:

$$\psi = A \cdot e^{2i\pi(p_x x + p_y y + p_z z - E t)/h} \tag{5′}$$

as the monochromatic plane wave that is associated with a current of light-quanta p_x, p_y, p_z, E.

The *intensity* of the wave is classically proportional to the square of the amplitude $A^2 = \psi \cdot \tilde{\psi}$ (~ means the transition to the complex conjugate). Now should that intensity be really transported by light-quanta then one would regard $\psi\tilde{\psi} = |\psi|^2$ as a measure of the *density* with which the light quanta are distributed in space. The number of light quanta in a volume dV is proportional to $\psi\tilde{\psi} \cdot dV$, and $\psi\tilde{\psi} \cdot dV$ shall be a measure of the *probability* of finding a light-quantum in dV. In that way, we have arrived at a modification of the **Bohr-Kramers-Slater** statistical picture. Namely, we can state, in connection with **Born**, **Jordan**, and **Dirac** ([1]): One can never obtain exact data about the microscopic distribution of light-quanta (nor ever observe them either, cf., the following no.), but only their mean distribution (for the observations, that would mean only data with a precision that mainly cannot be reduced, resp.). Calculating the probability function $\psi\tilde{\psi}$ will happen with the help of wave theory, but in which the wave state quantity ψ will take on no other meaning than just that of a *probability amplitude* whose square $|\psi|^2$ will give the probability for the appearance of corpuscles in reality. The ψ-field is only a "virtual guiding field" for the real light corpuscles.

Now, the fact that the probability amplitude ψ is complex is significant, i.e., that it possesses both magnitude and phase. The linear homogeneous oscillation equation has the property that the superposition of its various complex solutions ψ_1, ψ_2, … will again yield a solution $\psi = \psi_1 + \psi_2 +$ … However, its intensity is now:

$$|\psi|^2 = |\psi_1|^2 + |\psi_2|^2 + \ldots + (\psi_1\tilde{\psi}_2 + \tilde{\psi}_1\psi_2) + \ldots \tag{7}$$

([1]) **A. Einstein**, Berliner Ber. (1925), pp. 3; **Born-Heisenberg-Jordan**, Zeit. Phys. **35** (1926), pp. pp. 557, Chap. 4, § 3; **P. Jordan**, Die Naturwiss. **15** (1927), pp. 105; **M. Born**, Die Naturwiss. **15** (1927), pp. 238; **W. Heisenberg**, Zeit. Phys. **40** (1926), pp. 501; **P. Jordan**, Zeit. Phys. **40** (1926), pps. 661 and 809; **P. A. M. Dirac**, Proc. Roy. Soc. **113** (1926), pp. 621.

which differs from the sum of the individual intensities $|\psi_1|^2 + |\psi_2|^2 + \ldots$ since the mixed terms $(\psi_1\tilde{\psi}_2+\tilde{\psi}_1\psi_2) + \ldots$ have been added to them. One does not add the probabilities $|\psi_k|^2$ then, but one does superpose the complex amplitudes ψ_k . The simultaneous presence of two waves ψ_1 and ψ_2 with closely-coincident properties ($\lambda_1 \sim \lambda_2$, $\nu_1 \sim \nu_2$, $\mathfrak{s}_1 \sim \mathfrak{s}_2$), which will then lead to two currents of light-quanta with similar light quanta ($p_1 \sim p_2$, $E_1 \sim E_2$, $\mathfrak{s}_1 \sim \mathfrak{s}_2$), will then lead to a statistical density function for the light-quanta that is not simply the sum $|\psi_1|^2 + |\psi_2|^2$, but a density distribution that corresponds completely to the intensity distribution of the two waves ψ_1 and ψ_2 with mutual interference, so one will have an *interference of the probabilities* for the presence of corpuscles at the individual locations in the spacetime continuum. One will also find entirely analogous relationships in the theory of material particles. Their distribution can also be calculated statistically with the help of a probability amplitude ψ as a solution to an oscillation equation, whereby the various solutions are once more capable of interference.

In that way, we have already touched upon the question of the domain of validity of the *causal law* that any later state can be determined from the current state, or when regarded empirically, that if the current state is observed with sufficient (arbitrary, resp.) precision then its later state can be calculated with sufficient (arbitrary, resp.) precision and can be established by observation. With the picture of quantum theory, in that statement one already finds the antecedent statement that a precisely-observed initial state is not only unattainable in practice, also *intrinsically*. Namely, if the physical phenomenon is governed by only the wave equations for the probability functions then that will mean merely that the current probability state of the corpuscle is coupled with a later probability state by a law, but the microscopically-exact initial state will go to a final state according to the laws of physics that is nothing like the final state microscopically, either theoretically or experimentally. From the probable initial distribution of corpuscles, one can *always* assume only a *probable* distribution in the final state, and indeed with all the more dispersion when the initial final states are further separated from each other in spacetime. One can now seek the laws that are still hidden to us from among the enveloping statistical laws of the physical phenomena, namely, the ones that the individual corpuscles obey, which will first lead to those observed probabilistic laws *in the mean*.

However, one has not the slightest hint regarding those hidden laws experimentally, rather it seems that the action quantity h fixes an *intrinsic* limit to microscopic observability. Conversely: A principle of the indistinguishability of certain individual states first creates space for the exhibition of quantum theory, and that principle is anchored empirically to the extent that experiments give us no reason to seek a much more complicated theory of distinguishable individuals in place of quantum theory, which is relatively simple in form. At the same time, everyone is naturally free to believe in a hidden causality between individuals. Similarly, one can also believe that someday infinitely-fast signals will be found by which the theory of relativity will then be stripped of its intrinsic meaning.

17. Limits of optical resolution. – It shall now be shown that on the basis of elementary optical experiments, the existence of the **Planck** quantum of action h, as the lower limit of precision for the measurement of an effect, has an equivalent in wave theory in the existence of a lower limit for the optical and harmonic resolution of wave-trains. In so doing, that quantum-theoretic restriction on precision will prove to be something that is known and "explained" in optics, and the corresponding explanation in *mechanics* will be prefaced.

In the theory of waves, there is a completely well-defined, not further divisible, smallest unit, namely, **Laue**'s *elementary ray bundle*. A ray bundle ([1]) consists of a bundle of parallel cones of rays whose vertices all lie on a small area Δf ($= \Delta\xi \cdot \Delta\eta$) and all possess the same opening angle $\Delta\Omega$ ($\Delta\Omega = \Delta\alpha \cdot \Delta\beta$, while Δf shall be assumed to be perpendicular to the axis of the cone), and which belong to the frequency interval $\Delta\nu$ and possess the time duration Δt, moreover. An *elementary* ray bundle is one for which the interval Δ satisfies the relation:

$$\Delta\nu\,\Delta t \cdot \frac{\Delta\alpha\,\Delta\xi}{\lambda} \cdot \frac{\Delta\beta\,\Delta\eta}{\lambda} = 1 \cdot 1 \cdot 1\,. \tag{8}$$

In that way, each of the three factors shall be equal to 1 in their own right. The elementary ray bundle then indicates the lower limit on the distinguishability of differences in color, time, direction, and type of source for the radiation of waves, in the following sense:

$\Delta\nu \cdot \Delta t = 1$: In order to recognize the difference $\Delta\nu$ between two frequencies ν and $\nu + \Delta\nu$ in an oscillating structure, one needs to know at least the time $\Delta t \approx 1 / \Delta\nu$, namely, the duration of a temporal *beat*. Conversely, in order to compose an aggregate of oscillations over a time interval Δt from purely-harmonic (i.e., temporally unlimited) oscillations, one will need harmonic components from an interval $\Delta\nu \approx 1 / \Delta t$. One can also express that as follows: Over a given time interval Δt, the ν of a truncated oscillating structure is determined up to a quantity $\Delta\nu \approx 1 / \Delta t$, at best. For a given $\Delta\nu$, the *age* of the aggregate of oscillations, i.e., the time since its emission, is determined up to quantities $\Delta\nu \approx 1 / \Delta t$ at best, because it already possesses a time duration of Δt in its own right.

$\Delta\alpha \cdot \Delta\xi / \lambda = 1$: In order to recognize the distance $\Delta\xi$ between the source points for a monochromatic wave structure that is emitted from two source points and is perpendicular (almost perpendicular, resp.) to their connecting line, one needs a segment of arc along the ray with a magnitude of at least $\Delta\alpha \approx \lambda / \Delta\xi$, namely, the interval along that direction along which a spatial beat (e.g., bright and dark diffraction lines) plays out. On the other hand, in order to superpose a wave-aggregate of limits arc opening $\Delta\alpha$ from spherical waves, one needs the distance between the centers of the spherical waves to be a line segment $\Delta\xi \approx \lambda / \Delta\alpha$. One can also express that as follows: For a given distance $\Delta\xi$ from the source, the direction of the emitted ray is determined up to an angle $\Delta\alpha \approx \lambda / \Delta\xi$, at best. For a given set of directions $\Delta\alpha$, the center of the ray is determined up to a magnitude of $\Delta\xi \approx \lambda / \Delta\alpha$, at best.

Corresponding statements are true for $\Delta\beta \cdot \eta / \lambda = 1$.

In order to extend what was just considered to one or more centers that are emitting interfering spherical waves, we shall now consider the case of *plane waves*. The formal connection between (8) and the spherical wave Ansatz:

([1]) **M. von Laue**, Ann. Phys. (Leipzig) **44** (1914), pp. 1197. See this *Handbuch*, Bd. 20, Chap. 9, nos. **4**, *et seq.*

$$\psi = \frac{a}{\nu}\, e^{2i\pi r/\lambda}$$

consists of the fact that one sets second-order terms in:

$$\frac{r}{\lambda} = (r_0 + \Delta\alpha\,\Delta\xi + \Delta\beta\,\Delta\eta)\cdot\frac{1}{\lambda_0}\left(1 - \frac{\Delta\lambda}{\lambda_0}\right) ,$$

namely:

$$\Delta\alpha\cdot\frac{\Delta\xi}{\lambda} , \qquad \Delta\beta\cdot\frac{\Delta\eta}{\lambda} ,$$

equal to 1. If one proceeds correspondingly with the Ansatz for *plane waves* in the $\mathfrak{s}$-direction:

$$\psi = e^{2i\pi(x\mathfrak{s}_x + y\mathfrak{s}_y + z\mathfrak{s}_z)/\lambda} ,$$

and sets the second-order terms in:

$$\frac{x}{\lambda}\mathfrak{s}_x = (x_0 + \Delta x)\left(\frac{1}{\lambda_0} + \Delta\frac{1}{\lambda}\right)\mathfrak{s}_x$$

equal to 1 then one will get:

$$\Delta x\,\Delta\left(\frac{1}{\lambda}\right)\mathfrak{s}_x\cdot\Delta y\,\Delta\left(\frac{1}{\lambda}\right)\mathfrak{s}_y\cdot\Delta z\,\Delta\left(\frac{1}{\lambda}\right)\mathfrak{s}_z = 1\cdot 1\cdot 1 \tag{9}$$

as a counterpart to (8).

One can make statements in regard to (9) that correspond to what was said about (8), namely, for $\Delta\nu\cdot\Delta t = 1$, as above:

$\mathfrak{s}_x\cdot\Delta x\cdot\Delta\left(\frac{1}{\lambda}\right) = 1$. In order to recognize the difference $\Delta\lambda$ in an aggregate of plane waves that is composed of two different wavelengths and advances in the x-direction, the structure must possess a length of at least $\Delta x = 1 : \Delta\left(\frac{1}{\lambda}\right)$. In order to compose a plane-wave structure of length Δx out of purely-harmonic linearly-unbounded plane waves, one needs harmonic components from an interval of at least $\Delta\left(\frac{1}{\lambda}\right) = \frac{1}{\Delta x}$.

One can also express that in the form: For a given length Δx of a plane-wave aggregate, λ is determined up to a magnitude of $\Delta\left(\frac{1}{\lambda}\right) = \frac{1}{\Delta x}$, at best. For a given $\Delta\lambda$, the *position* of the wave aggregate is determined up to a magnitude of $\Delta x \approx 1 : \Delta\left(\frac{1}{\lambda}\right)$.

The experimental facts of wave optics that were just summarized shall now be translated into the language of *the theory of light-quanta* (¹). With the help of **Planck**'s and **de Broglie**'s equation:

$$E = h\,\nu\,, \qquad p_x = \frac{h}{\lambda}\,\mathfrak{s}_x\,, \qquad \ldots,$$

one will get from (8) and (9) that:

$$\Delta E\,\Delta t \cdot p\,\Delta\alpha\,\Delta\xi \cdot p\,\Delta\beta\,\Delta\eta = h \cdot h \cdot h\,, \tag{10}$$

$$\Delta p_x\,\Delta x \cdot \Delta p_y\,\Delta y \cdot \Delta p_z\,\Delta z = h \cdot h \cdot h\,. \tag{11}$$

That leads to the following statements about light-quanta as a consequence of the wave-optical theorems above:

$\Delta E\ \Delta t = h$. In order to establish the energy E of a light-quantum with a precision of ΔE, one needs a train of light-quanta of time duration $\Delta t = h / \Delta E$, such that in that way the age of each light-quantum (viz., the time t since it was emitted) will be undetermined up to Δt. By contrast, should the age of each light-quantum be known up to Δt (which is the case for only a truncated train of time duration $\leq \Delta t$) then the energy of each individual light-quantum would be undetermined up to $\Delta E = h / \Delta t$. Therefore, $\Delta E \cdot \Delta t = h$ is the smallest energy-time interval inside of which distinguishing the differences in energy and age of light-quanta will no longer be possible.

$p\ \Delta\alpha\ \Delta\xi = h$: In order to establish the source-point of light-quanta that propagate perpendicularly (almost-perpendicularly, resp.) to the x-direction with a precision of $\Delta\xi$, one will need a bundle of light-quanta with an angular opening of $\Delta\alpha = h / (p\ \Delta\xi)$, which will leave the direction of each light-quantum undetermined up to $\Delta\alpha$.

If one would like to know the direction of each light-quantum with a precision of $\Delta\alpha$, which is possible only in a bundle with an arc opening $\leq \Delta\alpha$, then the location of the origin of each individual light-quantum will be undetermined up to $\Delta\xi = h / (p\ \Delta\alpha)$. $p\ \Delta\alpha\ \Delta\xi = h$ is the smallest interval of directions and source locations inside of which making a distinction will no longer be possible.

$\Delta p_x\ \Delta x = h$: In order to establish the impulse p_x of light-quanta with a precision of Δp_x , one will need light-quanta that are distributed along a segment $\Delta x = h / \Delta p_x$. If one would like to know the location of a light-quantum up to Δx then one would need a truncated train of length $\leq \Delta x$. However, p_x will be determined only up to $\Delta p_x = h / \Delta x$ in it. $\Delta p_x \cdot \Delta x$ is the smallest region inside of which a distinction between the impulse and coordinate components of light-quanta will no longer be possible.

One can establish those statements by calculation in such a way that one will either completely abandon the idea of ascribing a well-defined value to E, t, x, y, x, p_x, p_y, p_z, ξ, η, α, β for a light-quantum and imagine that every light-quantum propagates over a region ΔE, Δt, … that

---

(¹) The theory of light-quanta for lattice diffraction was presented by **W. Duane**, Proc. Nat. Acad. America **9** (1923), pp. 158. Furthermore, see **G. Breit**, *ibidem*, **9** (1923), pp. 238; **A. H. Compton**, *ibidem*, **9** (1923), pp. 359; **P. S. Epstein** and **P. Ehrenfest**, *ibidem*, **10** (1924), pp. 133; *ibid.* **13** (1927), pp. 400.

corresponds to the conditions (10), (11), or one can further consider the light-quanta to be in and of themselves *points* with sharply-defined values of position, time, impulse, and energy, but whose magnitude can, in no case, be physically established more precisely than the limits of error $\Delta E\, \Delta t = h$, $\Delta p_x\, \Delta x = h$, etc. If one accepts the second picture, so one would like to imagine that light-quanta are absolutely sharp, then one would have to think that the ages of the individual light-quanta are distributed in a *statistically-disordered* fashion over the time domain $\Delta t = h / \Delta E$ for a value of E for a light-quantum train that is known up to ΔE. Corresponding statements are true for the regions $\Delta p_x \cdot \Delta x = h$, $p\, \Delta\alpha\, \Delta\xi = h$, $p\, \Delta\beta\, \Delta y = h$ as elementary domains. Since an exact localization of the light-quantum in the spaces of energy, time, impulse, and coordinates is basically excluded by that, the picture of a sharp quantum can be regarded as one that is only suited to many purposes, but for which there is no possibility of measuring it exactly.

18. Interference of matter. – The connection between the corpuscular theory and the wave theory of optics that was just described can now be carried over to mechanics directly and will then lead to a novel picture as the foundations of the *wave theory of matter* (^{1}). In classical mechanics, the energy E, the age t (time since emission), impulse p_x, p_y, p_z, position x, y, z, etc., of each mass-point are regarded as physically well-defined data that are basically observable with arbitrary precision. However, if one pictures the mass-point as a mechanical wave-field that is associated with the **Planck-de Broglie** equations $E = h\ \nu$, $p = h\ /\ \lambda$ then one will arrive at the equations for it:

$$\Delta E\, \Delta t \cdot p\, \Delta\alpha\, \Delta\xi \cdot p\, \Delta\beta\, \Delta\eta = h \cdot h \cdot h\,, \tag{12}$$

$$\Delta p_x\, \Delta x \cdot \Delta p_y\, \Delta y \cdot \Delta p_z\, \Delta z = h \cdot h \cdot h\,. \tag{13}$$

just as one does for light-corpuscles, which express the following theorems that assert the *interference of matter waves.*

$\Delta E \cdot \Delta t = h$. In order to establish the energy of freely-flowing mass-particles with a precision of ΔE, one needs an aggregate of such particles (extracted from an atomic beam) of temporal extent $\Delta t = h\ /\ \Delta E$, such that in that way the age t of an individual particle (i.e., the time since it was emitted from the atomic radiator) will be undetermined up to Δt. Conversely, should the age t of the particle be known with precision Δt (that would be the case for only a truncated particle ray of time duration $\leq \Delta t$) then the energy E of each individual particle would remain determinate only up to $\Delta E = h\ /\ \Delta t$. $\Delta E \cdot \Delta t = h$ is the smallest energy-time interval inside of which a more precise localization of the particle would no longer be possible.

$p\ \Delta\alpha\ \Delta\xi = h$: In order to establish the location of the source of particles that flow forth perpendicular to the x-direction with a precision of $\Delta\xi$, one will need a ray bundle with an angular opening of at least $\Delta\alpha = h\,/\,p\,\Delta\xi$, by which the direction of each particle will remain undetermined up to $\Delta\alpha$. Conversely, should the direction of each particle be known with a precision of $\Delta\alpha$

(^{1}) The fact that **Einstein**’s theory of gases can also be interpreted as a wave theory of matter, in the sense of **de Broglie**, was shown by **Schrödinger**, Phys. Zeit. **27** (1926), pp. 95.

(which is the case only for a ray bundle of opening $\leq \Delta\alpha$), then the location of the source of each individual particle could be determined only up to $\Delta\xi = h / p\, \Delta\alpha$. An exact localization of particles inside of a region $\Delta\alpha\, \Delta\xi = h / p$ would not be possible.

$\Delta p_x\, \Delta x = h$: In order to establish the impulse p_x of a particle with a precision of Δp_x, one will need a section of a particle ray of length $\Delta x = h / \Delta p_x$, and in that way, the location of each individual particle will remain undetermined up to Δx. Conversely, in order to know the position of a particle up to Δx, one will need a truncated ray of length $\leq \Delta x$. However, for such a thing, the impulse of the particle can be determined only up to $\Delta p_x = h / \Delta x$.

Those statements about the limits of "mechanical resolution" can also be expressed in a form that reproduces the interference of light rays when one associates the mass-points with *waves* with the help of the **Planck-de Broglie** equations:

$$E = h\,\nu, \qquad \mathfrak{p} = \mathfrak{s} \cdot \frac{h}{\lambda}$$

and postulates a spatial and temporal *interference* of those waves. The limit of resolution of those waves will be characterized by equations (8), (9):

$$\Delta\nu\,\Delta t \cdot \frac{\Delta\alpha\,\Delta\xi}{\lambda} \cdot \frac{\Delta\beta\,\Delta\eta}{\lambda} = 1 \cdot 1 \cdot 1\,,$$

$$\Delta x\,\Delta\left(\frac{\mathfrak{s}_x}{\lambda}\right) \cdot \Delta y\,\Delta\left(\frac{\mathfrak{s}_y}{\lambda}\right) \cdot \Delta z\,\Delta\left(\frac{\mathfrak{s}_z}{\lambda}\right) = 1 \cdot 1 \cdot 1\,,$$

whose derivation for light was explained thoroughly. The *interference of matter* then consists of the following:

$\Delta\nu\,\Delta t = 1$: If one lets two beams of atoms that include particles with the *two* values for kinetic energy E and $E + \Delta E$ fall on a screen then the amount of matter that falls upon it per unit time cannot be constant in time, but must possess a slow, regular beat period in time of duration $\Delta t = h / \Delta E$.

$\Delta\alpha\,\Delta\xi / h = 1$: If one lets particles with the same impulse p start from two source-points that are separated by $\Delta\xi$ then the *number* of particles per unit time that appear on a screen that is placed at a great distance cannot be constant for all angles α in a small region around $\alpha = 0$ (which is perpendicular to the x-direction), but must exhibit *maxima* and *minima* with the angular period of $\Delta\alpha = h / (p\,\Delta\xi)$.

$\Delta x\, \Delta\, (1 / \lambda) = 1$: In a stream that consists of particles with two different values of impulse p_x and $p_x + \Delta p_x$, the density of matter along the path x is not constant, but will possess a spatial period of $\Delta x = h / \Delta p_x$.

Those statements about the *interference of matter*, which are complete reproductions of the ones for the interference of light with the help of the equations of association $E = h\,\nu$, $\mathfrak{p} = \mathfrak{s}\, h / \lambda$, have still had no exact experimental confirmation in that special form up to now.

However, convincing direct proof of the wave nature of ponderable matter comes from the experiments in the reflection of electron beams by crystal lattices that **Davisson** performed, together with **Kunsman** and **German** (1). They dealt with discovering just those maxima of selective reflection in certain directions that might be like the ones that would appear under the reflection of Röntgen rays by the crystal in question. In that way, they confirmed the **de Broglie** relation $\lambda = h : p$, with the interpretation that was first given by **W. Elsasser** (2). The analogue of the **Debye-Scherrer** method was also examined: **Thomson** and **Reid**, as well as **E. Rupp** (3), found the diffraction rings that were well-known for Röntgen rays by passing a beam of electrons through thin metal foils. Finally, **Dymond** (4) found characteristic diffraction maxima for the reflection of electrons by helium gas atoms. It is also possible to interpret the **Ramsauer** effect as a diffraction phenomenon (5).

A conclusive indirect proof of the wave nature of matter is given by the experimentally-exact *confirmed results* that were indicated by the further development of wave mechanics for particles in force fields (e.g., electrons in the field of a nucleus or the field of light), which the next section will report on.

Due to the intrinsic impossibility of verifying the position, time, impulse, and energy data for a particle exactly, the depiction of a beam of atoms or electrons as something that consists of isolated individual corpuscles can be regarded as only an intuitive picture for many purposes that cannot be taken literally, since it does not take into account the very fact of its potential for interference. The same thing is then true for the pictures in the kinetic theory of gases (no. **21**) and the image of the atom being constructed from electric corpuscles.

19. Uncertainty relation. – According to **Heisenberg** (6), the impossibility of making simultaneous exact observations of the conjugate quantities E and t, p_x and x, etc., can already be seen from both the point-mechanical and point-optical standpoints. Suppose that one establishes the momentary position x of a moving electron. That can be accomplished by illuminating the electron and observing the scattered light. However, due to the wave nature of light, the position x of the emission of scattered light is fixed only up to an error $\Delta x \approx \lambda$ with the same order of magnitude as the wavelength. If one makes λ ever smaller in order to make that error smaller then under the scattering of light by an electron, the electron will suffer an even-greater change in impulse $\Delta p_x \approx h / \lambda$ (viz., the Compton effect). The smaller the limit of error Δx on the determination of position, the larger will be the limit of error Δp_x on the determination of the impulse of the electron, and calculation will then yield the connection $\Delta p_x \cdot \Delta x \approx h$. The same thing is true for light-quanta. That can happen when one sends an electron along its way and observes

---

(1) **Davisson** and **Kunsman**, Phys. Rev. **22** (1923), pp. 243. – **Davisson** and **German**, Nature **119** (1927), pp. 558; Phys. Rev. **30** (1927), pp. 705.

(2) **W. Elsasser**, Naturwiss. **13** (1925), pp. 711.

(3) **G. P. Thomson** and **A. Reid**, Nature **119** (1927), pp. 890; *ibid.*, **120** (1927), pp. 802. – **E. Rupp**, Ann. Phys. (Leipzig) **85** (1928), pp. 711.

(4) **E. G. Dymond**, Phys. Rev. **29** (1927), pp. 433.

(5) **L. Mensing**, Zeit. Phys. **45** (1927), pp. 603.

(6) **W. Heisenberg**, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," Zeit. Phys. **43** (1927), pps. 172 and 809; **N. Bohr**, Naturwiss. **16** (1928), pp. 245.

the location of the origin of an electron that is scattered by a light-quantum. However, due to the wave nature of matter, it can be established only up to an error of $\Delta x \approx \lambda$, in which $\lambda = h / p$ (p = electron impulse). Under electron scattering, the colliding light-quantum will suffer a change in impulse of $\Delta p_x \approx h / \lambda$ (Compton effect). The smaller the limit of error Δx in the determination of the position the light-quanta the larger will be the limit of error Δp_x in the determination of its impulse.

As **Heisenberg** remarked, one can consider the experimental impossibility of a simultaneous exact determination of the values of two canonically-conjugate quantities as the *prerequisite* for the presentation of quantum theory, just like the impossibility of sending signals faster than the speed of light is a prerequisite for the special theory of relativity. If there is an experiment that makes it possible to determine p and q sharply and simultaneously then that would contradict wave mechanics. The uncertainty relations:

$$\Delta p_x \cdot \Delta x \approx h \, ,$$
$$\Delta E \;\cdot \Delta t \approx h$$

are a counterpart to the "commutation relations" that were introduced in quantum mechanics (Chap. VII):

$$p_x \, x - x \, p_x = \frac{h}{2i\pi} \, , \qquad t \, E - E \, t = \frac{h}{2i\pi} \, .$$

One will carry that uncertainty in any calculation when one subdivides the entire region for which the values of E, t, ξ, α, etc., are available for light (matter, resp.) into elementary cells of volume (12), (13):

$$\Delta E \, \Delta t \cdot p \, \Delta\alpha \, \Delta\xi \cdot p \, \Delta\beta \, \Delta\eta = h^3 \, ,$$
$$\Delta p_x \, \Delta x \cdot \Delta p_y \, \Delta y \cdot \Delta p_z \, \Delta z = h^3 \, ,$$

which is the smallest volume inside of which E, t, ξ, α, etc., can no longer be distinguished. The number of elementary cells that are found in a finite region (e.g., cavity, color range) is then the *number of degrees of freedom* in that region. The elementary cell is identical to the **Laue** elementary ray-bundle (8) by way of $E = h\nu$, $p = h / \lambda$. The number of *degrees of freedom* in a finite region is equal to the number of light (matter, resp.) waves in the elementary ray-bundle that contains them.

20. Corpuscular and wave theory. – The corpuscular theory of light, as well as that of matter, is characterized by corpuscular *conservation laws* for mass, energy, and impulse of isolated indestructible and recognizable particles of light or matter whose history can be followed, while the wave theory is characterized by *interference* phenomena for which light + light = dark, matter + matter = vacuum will contradict the postulates of conservation.

For the empiricist, it suffices possess a theory that can predict the observed phenomena with sufficient precision in all of the cases that might occur and give the intrinsic limits of the precision to such predictions in any given case. That problem also shows up in wave mechanics, since it initially gives a recipe for calculating a certain wave amplitude ψ in space and time (which is complete, moreover, i.e., it contains a phase) that is rooted in an undulatory sphere of ideas. The square of its absolute value $\psi\tilde{\psi} = |\psi|^2$ will then be interpreted as the *intensity* of the light (the *density* of the matter, resp.) in the sense that $|\psi|^2$ is the *probability*, while ψ itself gives the "probability amplitude," along with the phase, that light (matter, resp.) corpuscles can be found in that spacetime cell and can influence the apparatus that is found there.

It is characteristic that such effects always appear in a corpuscular way, e.g., as the material precipitation of isolated atoms or as photochemical effects of isolated light-quanta, but not as wave-aggregates that use the prediction of the effect in an essential way. The legacy of the origin of the wave no longer lies in its *uncertainty*, which initially enters into the calculations as the *width* of a light or matter ray diffraction line, and which can predict the points of impact for the individual particles only roughly when it is interpreted in a corpuscular context. However, in the final analysis, one can [as **Darwin** suggested (^{1})] regard that reinterpretation of the *intensity-density* as a corpuscular probability as a concession to our habit of explaining observational data in terms of corpuscles. One can just as well keep to the familiar undulatory interpretation of light precisely. Here, one must envision that the present place of matter in history is the same as the place of light before the time of **Fresnel**.

We would like to develop **Darwin**'s concept somewhat further. If we consider, say, an electron current then from the corpuscular viewpoint, a certain velocity distribution (say, **Maxwell**'s) will be created at every location. By contrast, from the undulatory viewpoint, one has a superposition of matter waves of different wavelengths. However, it amounts to the same thing when one speaks of a particle current or a wave-train. That is because when an experimenter can verify and isolate individual particles with velocities that lie above or below the mean velocity, those individual velocities are by no means present already in the total beam, but one can say that on the contrary, the observation apparatus itself creates those individual homogeneous velocities. Likewise, the spectroscope first generates the individual colors from white light, and it is only an abuse of language when one says that the colors are already "included" in the white light, since that inclusion means nothing but a reference to the spectroscopic experiment. If one had believed in the old quantum theory that the **Stern-Gerlach** experiment proved the unique existence of individual atoms in varying stationary states in the magnetized beam of atoms then one would now prefer to say that the atomic beam represents a wave-train that is split into certain components by the apparatus in question, while other components might possibly be stripped off in a different apparatus that would be, however, less characteristic of the original atomic beam than they would be of the apparatus itself.

Empirically, there also exists an essential difference between the predictions of the old and the new conceptions of things in certain cases. Let us consider, say with **Darwin**, a modified **Stern-Gerlach** experiment. The atomic beam goes along the x-direction and passes through an inhomogeneous magnetic field that points parallel to the y-axis, in which it is split into a beam that is polarized along $+y$ and one that is polarized along $-y$. The $-y$ component will be screened out, while the $+y$ component will go through a *homogenous* magnetic field that is directed parallel to the z-axis, and in which it will be split into two components that are parallel to $+z$ and $-z$ according

---

(^{1}) **C. G. Darwin**, Proc. Roy. Soc. **117** (1927), pp. 268.

to the old theory. If they now further go through an inhomogeneous field that is parallel to y then they will again be directed along $\pm y$ and should ultimately give *two* lines on a screen behind it.

By contrast, according to wave mechanics, the z-field has the effect of rotating the plane of polarization of the rays that enter it and are polarized along $+ y$ around z as the axis of precession with the **Larmor** frequency. Therefore, if the homogeneous z-field has precisely the length that would be necessary for the atoms that enter it and are polarized along $+ y$ to suffer a precession around the z-direction by $n \cdot 180^{\circ}$ (n = whole number), such that they will enter into the last inhomogeneous y-field with a polarization that is only in precisely the $- y$ or $+ y$ direction, then they will no longer be split, and one should expect only *one line* on the screen.

21. Statistics of light quanta and gaseous atoms. – Define standing waves in a volume V with discretely-distributed wavelengths λ in such a way that the walls will be nodal surfaces. According to **Jeans** (1), the number of eigen-oscillations in the region from λ to $\lambda + \Delta\lambda$ is equal to [see Chap. 8, eq. (4)]:

$$\Delta Z = V \frac{4\pi}{\lambda^2} \Delta\left(\frac{1}{\lambda}\right) . \tag{14}$$

If one would like to employ the frequency $\nu = u : \lambda$, instead of λ, with u = phase velocity, then one would have:

$$\Delta\left(\frac{1}{\lambda}\right) = \Delta\left(\frac{r}{u}\right) = \frac{\Delta\nu}{u}\left(1 - \frac{\nu}{u}\frac{du}{d\nu}\right) = \frac{\Delta\nu}{g},$$

with the group velocity g, and one would then get:

$$\Delta Z = V \frac{4\pi\nu^2}{u^2 g} \Delta\nu . \tag{15}$$

The corpuscular interpretation of (1) with the help of the **de Broglie** association $p = h / \lambda$ gives the number of discrete impulse cells that lie in the range of impulse from p to $p + \Delta p$ as:

$$\Delta Z = V \frac{4\pi\, p^2}{h^3} \Delta p . \tag{16}$$

For *light*, those expressions are provided with a factor of 2 in order to account for the two degrees of freedom for polarization. With the corpuscular interpretation of (15), we can distinguish two special cases, the first of which is a light-quantum *in vacuo* (a non-dispersive medium) and the

---

(1) **I. H. Jeans**, Phil. Mag. **10** (1905), pp. 91. See also **M. Planck**, *Vorlesungen über die Theorie der Wärmestrahlung*.

second of which is that of mass-points μ with no mutual couplings and in the absence of an external force-field (i.e., an ideal gas). We will then have:

$$\varepsilon = h\,\nu\,,$$

and

$$u^2 \cdot g = c^3 \qquad \left(p = \frac{\varepsilon}{c}\right) \text{ for light,}$$

$$u^2 \cdot g = v^3 \qquad \left(\mu v = p = \sqrt{2\mu\varepsilon}\right) \text{ for a gas,}$$

so from (15):

$$\Delta Z = V\,\frac{8\pi\,\varepsilon^2}{c^3\,h^3}\,\Delta\varepsilon \text{ for light,} \tag{17}$$

$$\Delta Z = V\,\frac{2\pi\,(2\mu)^{3/2}}{h^3}\cdot\varepsilon^{1/2}\cdot\Delta\varepsilon \text{ for a gas} \tag{18}$$

will be the number of discrete energy cells that belong to the energy interval ε to $\varepsilon + \Delta\varepsilon$.

If one counts N eigen-oscillations in increasing order of ε from $\varepsilon = 0$ onwards and integrates (17), (18) then one will get:

$$Z(\varepsilon) = V\cdot\frac{8\pi\,\varepsilon^3}{3c^3\,h^3} \qquad \text{conversely} \qquad \varepsilon(Z) = \left(\frac{3Z}{8\pi V}\right)^{1/3}\cdot hc \quad \text{(light),} \tag{17′}$$

$$Z(\varepsilon) = V\,\frac{2\pi\,(2\mu\,\varepsilon)^{3/2}}{h^3\cdot 3/2} \qquad \text{conversely} \qquad \varepsilon(Z) = \left(\frac{3Z}{4\pi V}\right)^{2/3}\cdot\frac{h^2}{2\mu} \quad \text{(gas),} \tag{18′}$$

$Z(\varepsilon)$ is the number of the eigen-oscillation $\varepsilon = h\,\nu$, $\varepsilon(Z)$ is the energy $\varepsilon = h\,\nu$ of the Z^{th} eigen-oscillation. It depends upon *volume*.

$$\varepsilon(Z) = \text{const. } V^{-1/3} \quad \text{for light,} \qquad \varepsilon(Z) = \text{const. } V^{-2/3} \quad \text{for a gas.} \tag{18″}$$

We shall now ask what the *statistical distribution* of light or gas particles over the energy cells of volume V might be. Three particles a, b, c might be distributed over four cells (nos. 1, 2, 3, 4). One has the following twenty possibilities for the occupation of cells by particles:

Cell no. 1	3	0	0	0	2	2	2	1	0	0	1	0	0	1	0	0	1	1	1	0
″ ″ 2	0	3	0	0	1	0	0	2	2	2	0	1	0	0	1	0	1	1	0	1
″ ″ 3	0	0	3	0	0	1	0	0	1	0	2	2	2	0	0	1	1	0	1	1
″ ″ 4	0	0	0	3	0	0	1	0	0	1	0	0	1	2	2	2	0	1	1	1
	$p_3 = 1$ $p_0 = 3$				$p_2 = 1$, $p_1 = 1$, $p_0 = 1$												$p_1 = 3$, $p_0 = 1$			

In the first four of those distributions, one has one cell with three particles and three cells with zero particles in each case ($p_3 = 1$, $p_0 = 3$). For the next twelve distributions, one has $p_2 = 1$, $p_1 = 1$, $p_0 = 2$, and for the last four distributions, one has $p_1 = 3$, $p_0 = 1$.

In **Bose-Einstein** statistics (1), all of those twenty distributions are regarded as *equally probable*, so they are counted with the same weight. That assumption means that one "distributes the cells over the particle numbers" by giving each cell that contains 0, 1, 2, or 3 particles the same probability. By contrast, with the corpuscular interpretation, in which three individual distinguishable particles *a*, *b*, *c* are distributed over the four cells according to probability, one would like to assign a weight to the distribution $\left|\begin{matrix}2\\1\\0\\0\end{matrix}\right|$ that is four times as great as that of the distribution $\left|\begin{matrix}3\\0\\0\\0\end{matrix}\right|$, because the latter can be realized in *only one* way, since all three particles fall in the first cell $\left|\begin{matrix}abc\\0\\0\\0\end{matrix}\right|$, while the former can be realized in *four* ways $\left|\begin{array}{c|c|c|c} abc & abd & acd & bcd \\ d & c & b & a \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right|$. That is what one does in **Boltzman**'s statistics of the ideal gas molecule, which leads to **Maxwell**'s law of distribution in statistical equilibrium. If **Bose** statistics deviates from that by counting each of the twenty distributions above with equal weight then that is based upon the fact that one gives up the picture of *independent* individual corpuscles. However, based upon **Boltzmann**'s statistics, one will also come to approximately the same statistical energy distribution as the **Bose** statistics when one imagines that the particle that is found in a cell is endowed with a *phase* and their energies are not simply added, but their amplitudes are superposed like the superposition of oscillations (2).

We shall also mention **Fermi** statistics (3), which differ from **Bose** statistics by the fact that the only distributions that are allowed are the ones for which no cell possesses more than *one* particle. Thus, in the example above, according to **Fermi**, only the last four distributions are allowed, while the first sixteen distributions are counted with weight zero. In general, with **Fermi**, only those distributions are allowed for which $p_n = 0$ for $n = 2, 3, \ldots$, and only p_0 and p_1 are non-zero. One can go from the results of **Bose** statistics to those of **Fermi** statistics with no further discussion by specializing to the case of $p_n = 0$ for $n > 1$. **Bose** statistics is verified for light-quanta, while **Fermi**

---

(1) **S. Bose**, Zeit. Phys. **26** (1924), pp. 178. – **A. Einstein**, Berl. Ber. (1924), pp. 261; *ibid.* (1925), pp. 318. – **E. Schrödinger**, Phys. Zeit. **27** (1924), pp. 95.

(2) **A. Landé**, Zeit. Phys. **33** (1925), pp. 571.

(3) **E. Fermi**, "Zur Quantelung des idealen einatomigen Gases," Zeit. Phys. **36** (1926), pp. 902.

statistics is verified for material particles. The latter is a generalization of the **Pauli** principle ([1]) that no more than *one* electron can occupy a quantum state in an atom to an arbitrary system of *n* equivalent particles.

We shall not go into the development of the various statistical Ansätze for corpuscles of light and matter here. The result for light-quanta in thermodynamic equilibrium is the **Planck** law of radiation. For material particles, the departure from the classical **Boltzmann** statistics of the **Bose-Einstein** and **Fermi** statistics becomes especially remarkable at low temperatures ("degeneracy phenomena").

Their experimental confirmation is complicated by the fact that it is precisely at low temperatures that the **Van de Waals** molecular forces perturb the ideal behavior of gases. Thus, before one can develop a quantum theory of *nonideal* gases, an unambiguous experimental confirmation must be performed. The relationships for electrons, which are 2000 times lighter, are much more favorable, since, e.g., the conducting electrons in a piece of metal define a "gas" that degenerates into a molecular gas at much higher temperatures. **Pauli** ([2]) has adapted **Fermi** statistics to conducting electrons in particular in the context of the question of why conducting electrons in an external magnetic field do not give rise to a strong magnetization of the metal, although every electron possesses a magnetic moment of magnitude 1 magneton according to **Uhlenbeck** and **Goudsmit**. However, every cell can be occupied by not only one electron, but two of them, as long as their magnetic quantum numbers are different, i.e., when the magnetic axes of the two electrons point in opposite directions that are parallel and anti-parallel to the external field. The electron gas is already largely degenerate at ordinary temperatures, since most electrons crowd together in the cells of the least-possible number Z with the least-possible energy ε, such that they will be found in a cell with an oppositely-oriented electron, and in that way will be magnetically neutralized.

Sommerfeld ([3]) arrived at a comprehensive theory of conducting electrons by an advanced discussion of the properties of electron gases in external fields and a consideration of the gradients of temperature and density. There are formulas for the effects of **Thomson**, **Peltier**, **Volta**, and for the relationships between electrical and thermal conductors, etc., that go to the formulas for the classical (viz., **Lorentz-Drude**) theory of conducting electrons for high temperatures (low density, resp.), but for low temperatures and high densities, they exhibit characteristic degeneracies that are measured in the experiments, and in that way, confirm the **Fermi** statistics.

([1]) **W. Pauli**, Zeit. Phys. **31** (1925), pp. 765. Moreover, see **Ornstein** and **Kramers**, "Kinetische Herleitung des FERMIschen Verteilungsgesetzes," Zeit. Phys. **41** (1927), pp. 481; **W. Heitler**, "Freie Weglänge und Quantelung der Molekültranslation," Zeit. Phys. **44** (1927), pp. 161.

([2]) **W. Pauli**, "Über Gasentartung und Paramagnetismus," Zeit. Phys. **41** (1927), pp. 81.

([3]) **A. Sommerfeld**, "Elektrontheorie der Metalle auf Grund der FERMIschen Statistik," Zeit. Phys. **47** (1928), pp. 1; *ibid.* **47** (1928), pp. 43; Naturwiss. **15** (1927), pp. 825.

CHAPTER III

UNDULATORY MECHANICS OF CONSERVATIVE SYSTEMS

22. Optical-mechanical analogy. – In Chapter I, the close relationship between geometrical optics and classical mechanics was portrayed, which has its origin in the analogy between the *index of refraction:*

$$n\,(q) = \frac{c}{u} = \frac{\text{vacuum velocity}}{\text{phase velocity}}$$

of light, which depends upon the position coordinates q, and the position-dependent quantity:

$$n = c \cdot \frac{\sqrt{2m[E - U(q)]}}{E}, \tag{1}$$

for a system of mass-points with a potential energy function $U\,(q)$ and the kinetic energy $T = \frac{1}{2} m \sum\sum g^{KL}\, \dot{q}_K^2\, \dot{q}_L^2$. The constant energy parameter E corresponds to the parameter ν of color, which is thought to be constant and upon which the index of refraction depends. That analogy leads to the juxtaposition of curved light rays in a refracting field $n\,(q)$ and the mechanical trajectories in a potential field $U\,(q)$. Moreover, it leads to the juxtaposition of the *eikonal surfaces* $\varphi\,(q, t) = S\,(q) - c\,t$, which are perpendicular to the light rays and advance with a phase velocity of $u = c\,/\,n$, and the *action surfaces* $\varphi\,(q, t) = S\,(q) - c\,t$, which are "perpendicular" (in the non-Euclidian sense of the line element $\sqrt{\sum\sum g^{KL}\, dq_K\, dq_L}$) to the trajectories and which advance with the velocity:

$$u = \frac{c}{n} = c \cdot \frac{E}{\sqrt{2m[E - U(q)]}}. \tag{1′}$$

In that way, S is determined by a complete integral $S\,(q_1, q_2, \ldots, \alpha_1, \alpha_2, \ldots)$ to the eikonal equation:

$$(\text{grad } S)^2 = n^2 = \frac{c^2}{u^2}. \tag{2}$$

In geometrical optics, one further considers the wave functions:

$$\psi = e^{2i\pi\varphi/\kappa} = e^{2i\pi(S - ct)/\kappa} \qquad (\textit{optics})$$

with the temporal period $\nu = c\,/\,\kappa$.

One then determines κ in such a way that $\nu = c\,/\,\kappa$ is identical to the color parameter ν, which is definitive of n. In the mechanical case, where n depends upon the parameter E instead, there is initially no reason to fix the period ν by a special choice of $\kappa = c / \nu$ in a special way. Here, quantum theory prefers a certain choice: κ shall be chosen such that $\nu = c / \kappa$ has the same relationship to the energy parameter E as in the **Planck** relation $\nu = E / h$, so $\kappa = c\, h / E$, which will make the periodic function of mechanics assume the form:

$$\psi = e^{2i\pi E(S-ct)/hc} .$$

From now on, in contrast to Chapter I, we would now like to understand S to be usual action function (see no. **12**) in mechanics, so we shall now write S, instead of $S\, E / c$, as we have up to now, such that we will now have:

$$\psi = e^{2i\pi(S-Et)/h}, \qquad \nu = \frac{E}{h} \qquad (\textit{mechanics}). \tag{4}$$

In that way, S is determined from the equation:

$$(\text{grad } S)^2 = \frac{E^2}{h^2} \qquad \text{with} \quad u = \frac{E}{\sqrt{2m(E-U)}}, \tag{5}$$

which is known in the form (no. **12**):

$$\frac{1}{2m} (\text{grad } S)^2 + U(q) - E = 0 \tag{5′}$$

as the **Hamilton-Jacobi** partial differential equation (H. J. D). ψ in (4) will then represent the "**de Broglie** wave function." Due to (1′), the frequency and wavelength of the action wave are then:

$$\nu = \frac{E}{h}, \qquad \lambda = \frac{u}{\nu} = \frac{h}{\sqrt{2m(E-U)}} = \frac{h}{p}, \tag{6}$$

in which $p = \sqrt{2m(E-U)}$ represents the total impulse of the mechanical system at the location q for the total energy E, and also in the general case of an N-point system.

23. Fundamental equation of wave mechanics (^{1}). – The transition to the optical wave equation for ψ was carried out in no. **8**, in which the differential quotient $\partial S / \partial x$ in the eikonal

---

(^{1}) **E. Schrödinger**, "Quantisierung als Eigenwertproblem," Ann. Phys. (Leipzig) **79** (1926), pps. 361 and 489; *ibid.*, **80** (1926), pp. 437. *Abhandlungen zur Wellenmechanik*, Leipzig, 1927.

equation (grad $S)^2 - c^2/u^2 = 0$ was replaced with the operator $\frac{\kappa}{2i\pi}\frac{\partial}{\partial x}$, and it then took the form of the operator equation:

$$\left(\frac{\kappa}{2i\pi}\right)^2\left(\frac{\partial^2}{\partial x^2}+\cdots\right)-\frac{c^2}{u^2} = 0\,,$$

and when that was applied to a function ψ, it gave the equation:

$$\left(\frac{\kappa}{2i\pi}\right)^2\Delta\psi-\frac{c^2}{u^2}\psi = 0\cdot\psi\,,$$

which goes to the wave equation:

$$\Delta\psi+\frac{4\pi^2\nu^2}{u^2}\psi = 0\,. \tag{7}$$

With **Schrödinger**, the corresponding transition in mechanics shall now be carried out. One replaces:

$$\frac{\partial S}{\partial q_k} \quad \text{with} \quad \frac{h}{2i\pi}\frac{\partial}{\partial q_k}\,,\ \ldots \tag{8}$$

in equation (5) and applies the operator equation that one obtains to a function ψ. If we next consider the special case in which the system consists of only a single mass-point m, so (grad $S)^2$ possesses the form $\left(\frac{\partial S}{\partial x}\right)^2+\cdots$, then one will obtain the operator:

$$\left(\frac{h}{2i\pi}\right)^2\left(\frac{\partial^2}{\partial x^2}+\cdots\right)-\frac{E^2}{u^2} = 0\,,$$

and when it is applied to ψ, i.e.:

$$\left(\frac{h}{2i\pi}\right)^2\Delta\psi-\frac{E^2}{u^2}\psi = 0\cdot\psi\,,$$

or ultimately, with the use of $u = E/\sqrt{2m(E-U)}$:

$$\boxed{\Delta\psi+\frac{8\pi^2 m}{h^2}[E-U(q)]\cdot\psi = 0.} \tag{9}$$

That is **Schrödinger**’s oscillation equation for the function $\psi(x, y, z)$. The temporal period ν is then regarded as being determined from $\nu = E/h$. Therefore, if $\psi(x, y, z)$ is a spatial function that solves the last equation then:

$$\psi(x, y, z, t) = \psi(x, y, z)\cdot e^{2i\pi Et/h} \tag{9$'$}$$

will be the associated spacetime function. The classical motion of a mass-point with the energy E in the potential field $U(x, y, x)$ will be associated with a solution ψ of **Schrödinger**'s oscillation equation in space and time.

We now consider the general case of a mechanical system, whose classical energy possesses the form:

$$E = T + U = \tfrac{1}{2} m \sum\sum g^{LK} \dot{q}_L \, \dot{q}_L + U(q)$$

as a function of the N coordinates q_K and velocity, with the coefficients $g^{LK} = g^{KL}$ that depend upon q. When equation (5′) is written out in detail [see (28.b) in no. **9**], it will assume the form:

$$\frac{1}{2m}\sum\sum g_{LK}\frac{\partial S}{\partial q_L}\frac{\partial S}{\partial q_K} + U(q) - E = 0 \qquad \left[H\left(q, \frac{\partial S}{\partial q}\right) - E = 0\right] \quad (10)$$

here.

Schrödinger's prescription of replacing $\frac{\partial S}{\partial q_K}$ with $\frac{h}{2i\pi}\frac{\partial}{\partial q_K}$ will now lead to a certain ambiguity in the desired oscillation equation. For example, if g_{LK} depends upon position q then $g_{LK}\, p^L p^K$ will be identical to the commuted product $p^L g_{LK}\, p^K$. By contrast, $g_{LK}\frac{\partial}{\partial q_L}\frac{\partial}{\partial q_K}\psi$ is a different function from $\frac{\partial}{\partial q_L} g_{LK}\frac{\partial}{\partial q_K}\psi$, since g_{KL} is not differentiated with respect to q_L in the former, but it might be in the latter. In order to exclude that ambiguity from the outset, one must write the classical Hamiltonian function in a *symmetrized* form. For example, one can imagine, say, replacing $g_{LK}\, p^L p^K$ with the more symmetric form $\frac{1}{6}$ ($g_{LK}\, p^L p^K$ + five further permutations of the sequence of three factors), and reinterpreting things in terms of differential operators. According to **Schrödinger**, one can use the following symmetrization in place of that: If g^* is the determinant of the g^{LK} then one starts from the symmetrized form:

$$\frac{1}{2m}\frac{1}{\sqrt{g^*}}\sum_L\sum_K p^K \sqrt{g^*}\, g_{KL}\, p^L + U(q) - E = 0\,. \tag{11}$$

Namely, if one replaces p^K with $h / 2i\,\pi \cdot \partial / \partial q_K$ in that then that will give it the form:

$$\frac{1}{2m}\left(\frac{h}{2i\pi}\right)^2 \cdot \frac{1}{\sqrt{g^*}}\sum_L\sum_K \frac{\partial}{\partial q_K}\sqrt{g^*}\,\frac{\partial}{\partial q_L}\psi + (U - E)\psi = 0\,, \tag{11′}$$

which can be written simply:

$$\frac{1}{2m}\left(\frac{h}{2i\pi}\right)^2 \cdot \Delta\psi + (U - E)\psi = 0\,, \tag{11″}$$

according to no. **9**. One has once more obtained the fundamental equation (9) then, except that now Δ means the generalized **Laplace** operator in a multi-dimensional space, into which the non-Euclidian metric:

$$(ds)^2 = \sum_K \sum_L g^{KL}\, dq_K\, dq_L$$

has been introduced. The domain of validity of the simple fundamental equation (9) has been extended considerably in that way. In particular, if the g^{KL} are spatially constant then the same thing will be true for g^*, and one can take the unsymmetrized form (10) to be the starting point of the transition from classical to undulatory mechanics.

One now asks whether and under what conditions finite solutions $\psi(q)$ of the **Schrödinger** equation are present. As such conditions, the boundary conditions will appear *first of all*, say, $\psi(q)$ must vanish strongly enough at infinity such that $\int |\psi|^2\, dq_1 \cdots dq_F$ remains finite, or for a finite region of the q_K (q_K is, say, an angle that varies between 0 and 2π), one can demand periodicity (single-valuedness, resp.) of ψ. Now, for such "natural" boundary conditions that are given by the nature of the problem, there will generally exist only finite solutions $\psi(q)$ when the parameter E in the oscillation equation possesses certain distinguished values E_1, E_2, … (*eigenvalues*), which belong to certain functions ψ_1 (q), ψ_2 (q), … (*eigenfunctions*) as solutions. Now, the physical statement of undulatory mechanics that the *eigenvalue* E_n that makes the **Schrödinger** equation for a mechanical system $H(q, p) - E = 0$ soluble is identical to the *energy values* of the system in the distinguished quantum states. The determination of the energy values is the reduced to the *eigenvalue problem* for a linear homogeneous partial differential equation in an N-dimensional coordinate space for natural boundary conditions that is associated with a classical mechanical system.

The associated eigenfunction ψ_n (q), when provided with the periodic time factor $e^{2i\pi E_n t/h}$, will then represent an oscillation process in that coordinate space (q-space). The eigenfunction can also be interpreted as the expression for certain physically-observable properties of the mechanical system, as will be explained later. Moreover, in some special cases, the oscillation equation can be arranged to be such that, along with the discrete spectrum of eigenvalues E_1, E_2, …, it also possesses a continuous spectrum of E-values in the interval E_a to E_b, which belongs to a continuous family of eigenfunctions ψ_a to ψ_b. That case appears, e.g., for systems of atomic electrons, which indeed possess all energy values between E = ionization energy and $E = \infty$, in addition to the discrete quantum energies E_n.

24. Correspondence with the classical theory. – The formal relationship between classical and undulatory mechanics takes the form of the juxtaposition of the equation:

$$H(q, p) - E = 0 \quad \text{or} \quad H\left(q, \frac{\partial S}{\partial q}\right) - E = 0 \quad (\textbf{Hamilton-Jacobi}) \quad (12)$$

and the oscillation equation in the form (10′):

$$\left\{H\left(q, \frac{h}{2i\pi}\frac{\partial S}{\partial q}\right) - E, \psi\right\} = 0 \qquad (\textbf{Schrödinger}). \qquad (12')$$

One will see the relationship between the **Schrödinger** theory and the quasi-classical **de Broglie** theory with the action waves:

$$\psi = A \cdot e^{2i\pi S/h}, \qquad \text{where } (\operatorname{grad} S)^2 = \frac{E^2}{u^2} = 2m(E - U) \qquad (\textbf{de Broglie}), \qquad (13)$$

which are endowed with the period $\nu = E / h$, when one substitutes the solution Ansatz $\psi = A \cdot e^{2i\pi S/h}$ in (11″). The differential equation for $S(q)$:

$$\frac{h}{2i\pi}\Delta S + (\operatorname{grad} S)^2 - 2m(E - U) = 0 \qquad (\textbf{Schrödinger}) \qquad (13')$$

will then remain. Except for the first term, which is proportional to *h*, that is identical to the classical determining equation for *S* in (13). For $h = 0$, **Schrödinger**'s oscillation function will go to **de Broglie**'s.

According to **Jordan** (^{1}), one can also clarify the correspondence in the following way: The classical equation (12) for *S* can be written in the form:

$$\left\{H\left(q, \frac{\partial S}{\partial q}\right) - E, 1\right\} = 0 \qquad (14)$$

as an operation that is performed on the constant function 1, i.e., as multiplication by 1. One now considers the differential equation:

$$\left\{H\left(q, \frac{\partial S}{\partial q} + \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - E, 1\right\} = 0 \qquad (14')$$

or when written out in detail in symmetrized form (cf. **11**):

$$\left\{\frac{1}{2m}\sum_L\sum_K \frac{1}{\sqrt{g^*}}\left(\frac{\partial S}{\partial q} + \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)\sqrt{g^*}\, g_{KL}\left(\frac{\partial S}{\partial q} + \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + U - E, 1\right\} = 0 .$$

(^{1}) **P. Jordan**, Zeit. Phys. **40** (1927), pp. 809.

Since the differentiations give $\partial\, 1 / \partial q_L = 0$ when one does the calculations, the last equation will reduce to the form (cf., no. **9**):

$$\frac{1}{2m}\left[\frac{h}{2i\pi}\Delta S + (\mathrm{grad}\, S)^2\right] + U - E = 0\ ,$$

which is identical to (13′). However, that means that the **Schrödinger** equation can be reduced to the form (14′) for S by the Ansatz $\psi = e^{2i\pi S/h}$, which is the counterpart to equation (14) for the classical action function. (14′) goes to (14) when $h = 0$.

25. Equation of oscillation from a variational principle. – In the **Lagrangian** function L of the classical-mechanical system L = kinet. – pot. energy = $H\,(q, p) = 2U$, one formally replaces:

$$p^K \ = \frac{\partial S}{\partial q_K} \qquad \text{with} \qquad \frac{h}{2i\pi}\frac{1}{\psi}\frac{\partial \psi}{\partial q_K}$$

and solves the variational problem:

$$\left.\begin{aligned} J_1 &= \int \psi^2 \cdot L\left(q, \frac{h}{2i\pi}\frac{1}{\psi}\frac{\partial\psi}{\partial q}\right)\cdot dv = \text{extremum},\\ J_2 &= \int \psi^2 dv \text{ as an auxiliary condition.}\end{aligned}\right\} \tag{15}$$

The dv means the invariant volume element $dv = \sqrt{g^*}\, dq_1\, dq_2 \ldots dq_N$. With the introduction of a **Lagrange** factor E, (15) will be equivalent to the variational problem:

$$O = \delta J_1 + E\ \delta J_2 = -\ \delta \int F\,(\psi, \psi_q, q)\, dq_1 \cdots dq_N\ ,$$

with

$$F = \frac{h^2}{4\pi^2}\frac{1}{2m}\sum_K\sum_L g_{KL}\frac{\partial\psi}{\partial q_K}\frac{\partial\psi}{\partial q_L}\sqrt{g^*} + \psi^2 (U - E)\sqrt{g^*}\ ,$$

and the **Euler** variational equation:

$$\sum_K \frac{d}{dq_K}\left(\frac{\partial F}{\partial \psi_{q_K}}\right) - \frac{\partial F}{\partial \psi} = 0 \qquad \left(\psi_{q_K} \text{ means } \frac{\partial\psi}{\partial q_K}\right).$$

However, with the use of the function F that was just written down, that is identical to the **Schrödinger** equation in the form (11′), (11″). When one starts from the ordinary nonsymmetrized Lagrangian function, that variational process will then lead automatically to the **Schrödinger** equation:

$$\Delta\psi - \frac{8\pi^2 m}{h^2}\,(E - U)\,\psi = 0$$

with the correctly-symmetrized **Laplacian** operator Δ. The variational problem for complex and time-dependent ψ is given in no. **35**.

As one sees from the form of (11′), the **Schrödinger** equation has a *self-adjoint* character, i.e., it possesses a symmetric differential operator that one can also read from back to front as:

$$\sum_K \sum_L \frac{\partial}{\partial q_L}\sqrt{g^*}\,g_{KL}\,\frac{\partial}{\partial q_K}\psi$$

when one switches the sequence of differentiations. However, it is known that the eigenfunctions of a *self-adjoint differential equation* define an *orthogonal system:*

$$\int \psi_m(q)\psi_n(q)\,dv = 0 \qquad \text{for } m \neq n \qquad (dv = \sqrt{g^*}\,dq_1\,dq_2\cdots), \tag{16}$$

and indeed, the orthogonal system is *complete*, so every function a (q) that is orthogonal to ψ_m :

$$\int a(q)\cdot\psi_n(q)\,dv = 0$$

is a linear combination a (q) = $\sum c_n\cdot\psi_n(q)$ of eigenfunctions ψ_n that differ from ψ_m and have constant coefficients c_n . One can then *normalize* that eigenfunctions to 1, i.e., provide them with constant factors that make:

$$\int \psi_m^2(q)\,dv = 1\,, \tag{17}$$

if necessary. We shall not go into the special relationships that relate to a continuous spectrum of eigenvalues.

The equivalence of the Schrödinger equation with a variational problem that likewise guarantees the orthogonality and completeness of the system of eigenfunctions is also significant on the grounds that a transition to *other coordinates* is much easier to accomplish with the integrand of the variational problem than it is with the differential equation itself. Furthermore, in the search for *approximate* eigenfunctions, it is useful to be able to apply the much-faster-converging approximation methods to the solution of the variation problem (e.g., the **Ritz** method), instead of applying the slowly-converging perturbation methods to the solution of the differential equation.

26. Rotator with an axis fixed in space. – As an example ([1]), we shall consider a rigid body that can rotate around a fixed axis. Its single degree of freedom is the angle of rotation φ. If its moment of inertia is J and $p = J \cdot \dot{\varphi}$ is its angular impulse then the classical total energy will be:

$$T + U = \frac{1}{2J}\, p^2 = E \qquad (U = 0),$$

and the wave equation will become:

$$\frac{d^2\psi}{d\varphi^2} - \frac{8\pi^2 J}{h^2} E\psi = 0$$

when uses the conversion (8). Its solution is:

$$\psi = A \cdot \sin\left(\sqrt{\frac{8\pi^2 E\, J}{h^2}} \cdot \varphi - \delta\right) .$$

Here, the *factor* of φ must be a *whole number n* in order for ψ to be a single-valued and continuous function of the angle of rotation $\psi(\varphi) = \psi(\varphi \pm 2\pi)$. That will give the condition:

$$\sqrt{\frac{8\pi^2 E\, J}{h^2}} = n , \qquad \text{i.e.,} \qquad E_n = \frac{n^2\, h^2}{8\pi^2 J}$$

for the quantum condition for the energy of the rotator. The eigenfunctions will then read simply:

$$\psi_n(\varphi) = \frac{1}{\sqrt{\pi}} \cdot \sin(n\,\varphi - \delta_n) .$$

They are normalized to 1 by way of the factor $1/\sqrt{\pi}$.

27. Rigid rotator with a free axis. – When one introduces the polar coordinates ϑ and φ, the impulses p_ϑ and p_φ, and the moment of inertia J, the classical energy will read:

$$T + U = \frac{1}{2J}\left(p_\vartheta^2 + \frac{p_\varphi^2}{\sin^2\vartheta}\right) = E \qquad (U = 0),$$

so the wave equation (9), with Δ in polar coordinates ϑ, φ will be:

([1]) For the following examples, see **E. Schrödinger**, Ann. Phys. (Leipzig) **79** (1926), pp. 489.

$$\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right)+\frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\varphi^2}+\frac{8\pi^2 J\,E}{h^2}\psi\ =0\ .$$

That is the differential equation of the *spherical functions*. In order for ψ to be single-valued and continuous on the sphere, the factor of ψ must be:

$$\frac{8\pi^2 J\,E}{h^2}\ = n\,(n+1), \qquad\qquad n = 0, 1, 2, \ldots$$

The energy levels are:

$$E_n = \frac{n(n+1)\,h^2}{8\pi^2 J}\ = \left[\left(n+\tfrac{1}{2}\right)^2-\tfrac{1}{4}\right]\cdot\frac{h^2}{8\pi^2 J}\ ,$$

accordingly, i.e., a formula with “half-integer” quanta, as would occur in the terms of *band spectra*. The associated eigenfunctions are the spherical functions:

$$\psi_n = Y_n\,(\varphi,\ \vartheta) = \sum_{m=0}^{n}(A_m\cos m\varphi+B_m\sin m\varphi)\sin^m\vartheta\frac{d^m P_n(\cos\vartheta)}{d\,(\cos\vartheta)^m},$$

which can be normalized to 1.

28. Harmonic oscillator. – For the *linear harmonic oscillator*, the classical energy function reads:

$$T+U = \frac{1}{2m}p^2+\frac{(2\pi\nu_0)^2 m}{2}q^2\ = E\ ,$$

if q means the distance from the mass-point to the rest position, ν_0 means the eigenfrequency, and $p = m\dot{q}$ means the impulse. The wave equation (9) in q-space is then:

$$\frac{d^2\psi}{dq^2}+\frac{8\pi^2 m}{h^2}\left(E-\frac{(2\pi\nu_0)^2 m}{2}q^2\right)\cdot\psi\ = 0\ .$$

With the help of the abbreviation:

$$x = q\cdot 2\pi\sqrt{m\nu_0/h}\,, \tag{18}$$

one will get the equation:

$$\frac{d^2\psi}{dx^2}+\left(\frac{2E}{h\nu_0}-x^2\right)\cdot\psi\ = 0 \tag{18′}$$

for ψ as a function of the dimensionless quantity x.

Now, the mathematical theory of that equation says that it is soluble for finite and continuous functions ψ only for special values of $2E / h\,\nu_0$, namely, for the *eigenvalues:*

$$\frac{2E}{h\nu_0} = 1, 3, 5, \ldots, 2n + 1, \ldots$$

The n^{th} eigenfunction $\psi_n(x)$ that belongs to the n^{th} eigenvalue:

$$E_n = \frac{h\nu_0}{2}\,(2n + 1), \qquad \nu_0 = \frac{E_n}{h}, \tag{18″}$$

when one affixes the time factor $e^{2i\pi\nu_n t}$, is equal to:

$$\psi_n(x) = H_n(x) \cdot e^{2i\pi t E_n / h}. \tag{19}$$

The $H_n(x)$ in that means the "n^{th} normalized **Hermite** orthogonal function":

$$\left.\begin{aligned} H_n(x) &= \frac{(-1)^n}{\sqrt{2^n \cdot n!}} \frac{d^n e^{-x^2}}{dx^n} \\ &= \frac{e^{-x^2}}{\sqrt{2^n \cdot n!}} \left\{ (2x)^n - \frac{n(n-1)}{1!}(2x)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!}(2x)^{n-4} - + \right\}. \end{aligned}\right\} \tag{19′}$$

The $H_n(x)$ are mutually-orthogonal and normalized to 1:

$$\int_{-\infty}^{+\infty} H_n(x)\,H_m(x)\,dx = \begin{cases} 0 & \text{for} \quad n \neq m, \\ 1 & \text{for} \quad n = m. \end{cases}$$

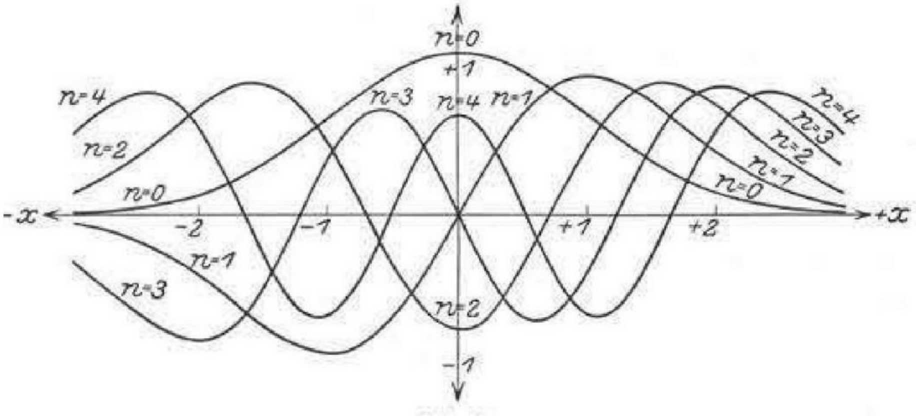

Figure 4.

The first five functions $H_0(x)$ to $H_4(x)$ are represented graphically in Fig. 4. It shows the similarity with the picture of the fundamental frequency and the harmonics of a string, and indeed and infinitely-long string with an inhomogeneous mass distribution that belongs to a position-dependent phase velocity $u = E/\sqrt{2m(E-U)}$ along the q-axis, which corresponds to the position-dependent expression for the potential energy $U(q)$ of the oscillators. The eigenfunctions are standing waves. The zeroes (i.e., nodes) of successive $H_n(x)$ are separate from each other. Except for $x = 3$, all five functions approach the x-axis monotonically. $H_n(x)$ has n finite zeroes and two zeroes at $\pm\infty$, as well as $n + 1$ antinodes.

The eigenvalues E_n in (18″) are given by *odd* multiples of $h\nu_0 / 2$. One will then get "half-integer" quantum levels, in contrast to the whole-number levels $E_n = n \cdot h\nu_0$, that establish the point-mechanical oscillator frequencies by quantum conditions. Experiments, especially with anharmonically oscillating systems, had already suggested the introduction of a special "zero-point energy" of magnitude $h\nu_0 / 2$ that would overlap with the whole-number quantized energy values. The new quantum theory does not require such an additional assumption. It also justifies the "half quantum numbers" that are often required by experiment in other cases.

29. Hydrogen atom (1). – The classical energy function, in polar coordinates r, ϑ, φ, reads:

$$T + U = \frac{1}{2\mu}\left(p_r^2 + \frac{1}{r^2}p_\vartheta^2 + \frac{1}{r^2\sin^2\vartheta}p_\varphi^2\right) - \frac{\varepsilon^2}{r} = E$$

here, and the Schrödinger equation (9), with the Δ in polar coordinates, will read:

$$0 = \frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{1}{r^2}\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right) + \frac{1}{r^2\sin^2\vartheta}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{8\pi^2\mu}{h^2}\left(E + \frac{\varepsilon^2}{r^2}\right)\psi \ .$$

We try to solve it by the Ansatz:

$$\psi(r, \vartheta, \varphi) = X(r) \cdot Y(\vartheta, \varphi)$$

and find, by substitution, that:

$$0 = Y(\vartheta,\varphi)\cdot\left[\frac{\partial^2 X}{\partial r^2} + \frac{2}{r}\frac{\partial X}{\partial r}\right] + \frac{X}{r^2}\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\varphi^2}\right] + \frac{8\pi^2\mu}{h^2}\left(E + \frac{\varepsilon^2}{r^2}\right)X\cdot Y \ .$$

That equation splits into two equations. Namely, if one sets the second square bracket equal to $-l(l+1)\cdot Y$:

---

(1) **E. Schrödinger**, Ann. Phys. (Leipzig) **79** (1926), pp. 361. Further, see, **J. Waller**, Zeit. Phys. **38** (1926), pp. 635; **C. Eckart**, Phys. Rev. **28** (1926), pp. 927.

$$\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial Y}{\partial\vartheta}\right)+\frac{1}{\sin^2\vartheta}\frac{\partial^2 Y}{\partial\varphi^2}+l\,(l+1)\,Y\ = 0 \qquad (l = 1, 2, \ldots)$$

then the equation for $X\,(r)$ will remain:

$$\frac{\partial^2 X}{\partial r^2}+\frac{2}{r}\frac{\partial X}{\partial r}+X\cdot\left(\frac{8\pi^2\mu}{h^2}E+\frac{8\pi^2\mu}{h^2}\frac{\varepsilon^2}{r}-\frac{l\,(l+1)}{r^2}\right) = 0\ .$$

One has the equation of the spherical functions with the eigen-solutions $Y = Y_l\ (\vartheta,\ \varphi)$ for Y. Investigation of the remaining differential equation for $X\,(r)$:

$$\frac{\partial^2 X}{\partial r^2}+\frac{2}{r}\frac{\partial X}{\partial r}+X\cdot\left(a+\frac{b}{r}-\frac{l\,(l+1)}{r^2}\right) = 0$$

will show that for every positive a, so for every *positive energy* E, it possesses solutions that are finite and continuous in the entire r-domain and converge to zero like $1/r$ at infinity. The eigenvalue (i.e., energy) spectrum is then *continuous* for positive E.

For *negative* E, one sees that it possesses solutions $X\,(r)$ that are finite and continuous and converge to zero at infinity sufficiently fast only when:

$$\frac{b}{2\sqrt{-a}} = n\ , \quad \text{i.e.,} \qquad \frac{2\pi\,\mu\,\varepsilon^2}{h\sqrt{-2\mu\,E}} = n, \qquad \text{and therefore} \quad n > l\ .$$

The energy value is then:

$$E_n = -\frac{2\pi^2\mu\,\varepsilon^4}{h^2\,n^2} \qquad (\textbf{Bohr}\text{'s energy values})$$

and that will yield the associated eigenfunctions as:

$$X_{nl}\,(r) = f_{nl}\,(x) =\ x^l\,e^{-x}\sum_{k=1}^{n-l-1}\frac{(-2x)^k}{k!}\binom{n+l}{n-l-1-k}\ ,$$

with the argument:

$$x =\ r\,\frac{2\pi\sqrt{-2\mu E}}{h}\ = \frac{r}{a_1\,n}\ , \quad \text{with} \quad a_1 = \frac{h^2}{4\pi^2\mu\,\varepsilon^2}\ ,$$

in which a_1 is radius of the first **Bohr** orbit for hydrogen. $f_{nl}\,(x)$ is known to be the $(2l + 1)^{\text{th}}$ derivative of the $(l + n)^{\text{th}}$ **Laguerre** polynomial. The eigenfunctions:

$$\psi_{nl}\,(r,\ \varphi,\ \vartheta) = \Psi_n\,(\varphi,\ \vartheta)\cdot X_{nl}\,(r)$$

can be normalized to 1. Fig. 5, which is taken from a paper by **L. Pauling**, shows the shape of the part of the ψ functions that depends upon r for various values of n and l. The number of nodes (zeros of ψ) is characteristic. It is equal n when one counts the zero at $r = \infty$. The symbols K, L, M are the Röntgen symbols that belong to n, l. Fig. 6 shows the mean charge density distribution (see, no. **30**) on the sphere of radius r, as measured by ψ^2. The vertical line divides the total charge into equally-large external and internal parts. The horizontal thick line along the abscissa axis shows the range of the electron orbit in the **Bohr** theory. For $l > 0$, the radial density function in the figure must be multiplied by a spatial angle function.

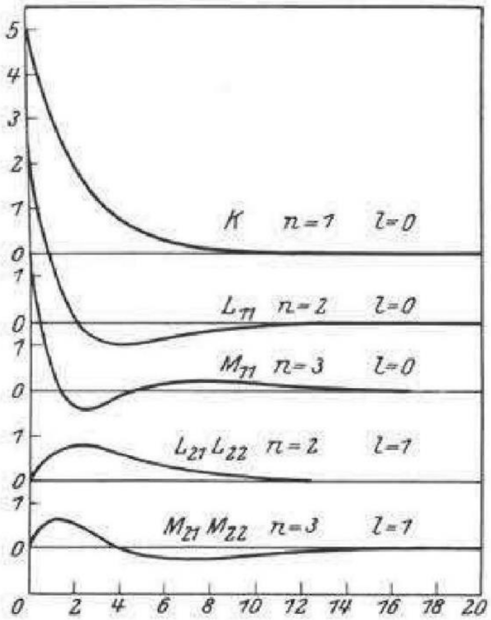

Figure 5.

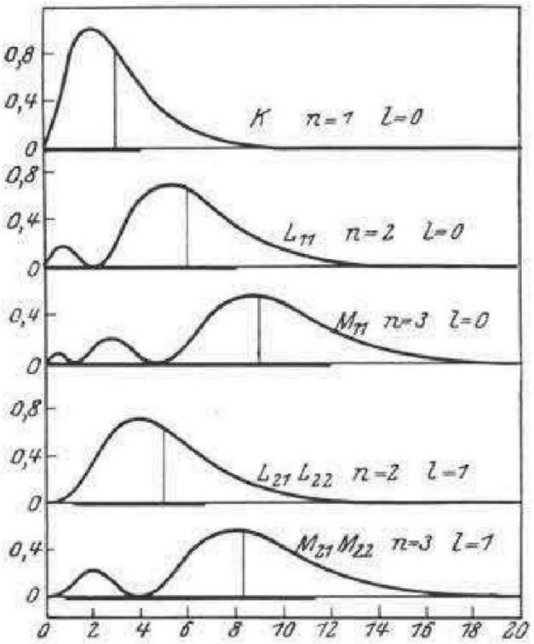

Figure 6.

We must pass over a discussion of the examples of system with more than one electron here. The examination of that problem has made it possible to explain a large number of questions in atomic physics that are connected with the spectra of higher atoms.

30. Continuous interpretation of the field scalar. – Whereas the point-mechanical theory of an N-body system admits a continuous family of orbits of the N mass-points in three-dimensional space, from which the older quantum theory selected a discrete number of "quantum orbits" with distinguished energy values E_1, E_2, … by a certain prescription, wave mechanics describes the state with the energy value E_k by an eigenfunction ψ_k in a space of $3N$ coordinates. In order to arrive at an interpretation of the ψ-function, we next consider the special case of a single mass-point in a temporally-constant potential field (e.g., the hydrogen electron in the field of a nucleus). The state of the system with an energy E_k is described by a scalar function ψ_k here that occupies all of three-dimensional space, instead of the k^{th} quantum orbit of **Bohr**'s theory. If the eigenfunctions are normalized to 1, i.e., $\int \psi_k^2\, dv = 1$, then we can imagine that the total energy E_k of the state in all of space combined has the form:

$$E_k = \int \psi_k^2(q)\, E_k\, dv\,, \qquad \int \psi_k^2(q)\, dv = 1\,,$$

in such a way that the amount of energy in the volume element dv is equal to $\psi_k^2\, E_k\, dv$, so the amount per unit volume at the location q is equal to $\psi_k^2\, E_k$. ψ_k^2 enters into that as a type of *density function* at the location q. One then arrives at the picture of the mass-point as not being restricted to the discrete **Bohr** quantum orbit, but being spread over all of coordinate space. For example, if ε is the electric charge of the mass-point then a corresponding *charge density* $\rho_\varepsilon(q) = \psi_k^2 \cdot \varepsilon$ will be created in the quantum state k, and a *mass density* $\rho_\mu(q) = \psi_k^2 \cdot \mu$, if μ is the mass of the point. One thus arrives a picture in which space is continuously occupied with mass and charge density, which deviates from the discrete quantum orbits of **Bohr**'s theory. Now, the calculation of that distribution of density in special cases shows that they mainly originate in those regions of space where the **Bohr** quantum orbits also run (see, e.g., Fig. 6). The continuous charge distribution will also make itself felt externally then as the effect of other charged particles with forces that do not deviate very much from those of the **Bohr** orbits ([1]).

If one has a system of several electrons, when expressed in point-mechanical terminology, then ψ_k will be a function in $3N$-dimensional space, and the same thing will be true of the density $\psi_k^2(q_1, q_2, \ldots, q_{3N}) = \psi_k^2(\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N)$. The charge density at a location $\mathfrak{r}$ in ordinary three-dimensional space is then additively composed of the densities that each individual charge would generate there, namely:

$$\rho_\varepsilon(\mathfrak{r}) = \varepsilon \sum_{L=1}^{N} \int \psi_k^2(\mathfrak{r}_1, \ldots, \mathfrak{r}_{L-1}, \mathfrak{r}, \mathfrak{r}_{L+1}, \ldots, \mathfrak{r}_N) \cdot dv_1 \cdots dv_{L-1}\, dv_{L+1} \cdots dv_N \tag{20}$$

(Each $\mathfrak{r}$ shall represent three coordinates, and similarly for dv.)

Moreover, in no. **16**, we became acquainted with a *statistical* interpretation for the ψ-function, instead of the continuous interpretation, and gave preference to it.

One will obtain the mean value of any point-mechanically defined physical quantity in the system, e.g., the mean moment in the state k, when one provides the function $\mathfrak{M} = \mathfrak{M}(\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N)$, which gives the moment of N point-charges in the configuration $\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N$, with the density factor $\psi_k^2(\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N)$, and integrates over the entire coordinate volume $dv = dv_1 \ldots dv_N$:

$$\mathfrak{M} = \int \mathfrak{M}(\mathfrak{r}_1, \ldots, \mathfrak{r}_{L-1}, \mathfrak{r}, \mathfrak{r}_{L+1}, \ldots, \mathfrak{r}_N) \cdot dv_1 \cdots dv_{L-1}\, dv_{L+1} \cdots dv_N\,. \tag{20′}$$

([1]) **A. Unsöld**, Ann. Phys. (Leipzig) **82** (1927), pp. 355.

Electrodynamically, radiation will emanate from an electric charge distribution $\rho_\varepsilon(q)$ in space only when the charge distribution (its moment, resp.) varies in time. Now, $\rho_\varepsilon = \varepsilon \cdot \psi_k^2(q)$ is the charge distribution on the state *k*, which depends upon only the configuration, but not time. That explains the fact that the stationary quantum states do not radiate. As is known, that was physically quite hard to understand in **Bohr**'s theory of the quantum orbits of moving electrons and had to be added to it as a special postulate.

If one were to now understand ψ_k to mean, not the function of position above, but the eigenfunction, when provided with the factor $e^{2i\pi E_k t/h}$ that is periodic in time, then the results above would still be true when one now writes $\psi_k \cdot \tilde{\psi}_k$ everywhere, in place of ψ_k^2, in which $\tilde{\psi}$ gives the complex conjugate to ψ. That is because in $\psi\tilde{\psi}$, one extends that time factor with its conjugate to $e^0 = 1$, and one will then get the same time-independent values for ρ, $\bar{\mathfrak{M}}$, etc., as above.

If we assume a result of the following section then we shall state at this point that a "transition density" is definitive of the transition from the state *k* to a state *l* :

$$\rho^{lk} = \psi_k(q)\psi_k(q) \cdot e^{2i\pi(E_k - E_l)t/h} = \text{amplitude} \cdot \text{period} \cdot \text{time} \cdot \text{factor}. \tag{21}$$

One notes that this "transition density" possesses the temporal period:

$$\nu = \frac{1}{h} \cdot (E_k - E_l) \tag{21′}$$

which *agrees with the* ***Bohr*** *frequency condition*. The combination principle then says that the periodic eigenfunction ψ_k of the initial state with $\nu_k = E_k / h$ and the periodic ψ_l of the final state with $\nu_l = E_l / h$, when combined, will create the *beat frequency* $\nu = \nu_k - \nu_l$. The radiation will vanish when the transition density (21) possesses zero amplitude everywhere, or in the many-body problem, when the density that is defined on the model of (21):

$$\rho^{lk} = e^{2i\pi(E_k - E_l)t/h} \sum_{L=1}^{N} \int \psi_k(\mathfrak{r}_1, \mathfrak{r}, \mathfrak{r}_N) \cdot \psi_l(\mathfrak{r}_1, \mathfrak{r}, \mathfrak{r}_N)\, dv_1 \cdots dv_{L-1}\, dv_{L+1} \cdots dv_N \tag{21″}$$

possesses zero amplitude. If that case occurs for two special eigenfunctions ψ_l, ψ_k then no radiative transition can exist between the two states *l* and *k*, so the transition in question is *forbidden*. That will be applied in no. **32**.

31. Eigenfunctions in the several-body problem (1). – We consider a system of *N* mass-points that do not perturb each other, since the total potential energy *U* is constructed from the

---

(1) **W. Heisenberg**, "Mehrkörperproblem und Resonanz in der Quantenmechanik," Zeit. Phys. **38** (1926), pp. 411; *ibid.*, **39** (1926), pp. 499; *ibid.*, **41** (1927), pp. 239.

potential energies $U^{(L)}$ of the individual particles by themselves. The **Schrödinger** oscillation equation in 3N-dimensional space for $\psi(\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N)$ will then read:

$$\sum \Delta_L \psi + \frac{8\pi^2\mu}{h^3}\left[E - \sum U^{(L)}(\mathfrak{r}_L)\right] \cdot \psi = 0\,, \tag{22}$$

in which Δ_L means the **Laplace** operator that arises from the kinetic energy of the L^{th} mass-point. One gets the solutions ψ here from the product Ansatz with the N factors:

$$\psi(\mathfrak{r}_1, \ldots, \mathfrak{r}_N) = \psi^{(1)}(\mathfrak{r}_1) \cdot \psi^{(2)}(\mathfrak{r}_1) \cdots \psi^{(N)}(\mathfrak{r}_N)\,, \tag{23}$$

which reduces (22) to the N equations ($L = 1, 2, \ldots, N$):

$$\Delta_L \psi^L + \frac{8\pi^2\mu}{h^3}\left[E^{(L)} - U^{(L)}(\mathfrak{r}_L)\right]\psi^{(L)} = 0 \qquad \text{with} \qquad E^{(1)} + \cdots + E^{(N)} = E\,. \tag{23′}$$

Each of the individual equations has an infinite system of eigenvalues and eigenfunctions:

$$E_1^{(L)},\ E_2^{(L)}, \ldots \text{ with } \quad \psi_1^{(L)},\ \psi_2^{(L)}, \ldots, \qquad \text{and in general} \qquad E_k^{(L)} \quad \text{with} \quad \psi_k^{(L)}(\mathfrak{r}_L)\,,$$

and the original equation (22) will then possess the product solutions:

$$\psi_{kl\ldots n} = \psi_k^{(1)} \cdot \psi_l^{(2)} \cdots \psi_n^{(N)} \qquad \text{with} \qquad E_{kl\ldots n} = E_k^{(1)} + E_l^{(2)} + \cdots + E_n^{(N)}\,, \tag{24}$$

in which $k, l, \ldots, n$ are any numbers that characterize the individual solution factors in ψ.

We would now like to assume, in particular, that all N mutually non-interacting mass-points move in *the same* potential field $U^{(1)}(\mathfrak{r}) = U^{(2)}(\mathfrak{r}) = \ldots = U^{(N)}(\mathfrak{r})$, say, all of them are in a field that depends upon only position, e.g., gas atoms in a closed vessel or mutually non-interacting electrons around a positive nucleus. The upper indices can then be dropped, because the series of eigenfunctions $\psi_1^{(K)}, \psi_2^{(K)}, \ldots$ will be identical to the series of eigenfunctions $\psi_1^{(L)}, \psi_2^{(L)}, \ldots$, for which we would then like to write simply $\psi_1, \psi_2, \ldots$ The same thing is true of the eigenvalues, which define only the simply-infinite series $E_1, E_2, \ldots$ The wave equation of the N-body problem of N mutually non-interacting in the same potential field will then possess a solution:

$$\psi^{(1)}(\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N) = \psi_k(\mathfrak{r}_1) \cdot \psi_l(\mathfrak{r}_2) \ldots \psi_n(\mathfrak{r}_N)\,, \qquad \text{with} \quad E = E_k + E_l + \ldots + E_n\,. \tag{25}$$

Similarly:

$$\psi^{(1)}(\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_N) = \psi_{k'}(\mathfrak{r}_1) \cdot \psi_{l'}(\mathfrak{r}_2) \ldots \psi_{n'}(\mathfrak{r}_N)\,, \qquad \text{with} \quad E' = E_{k'} + E_{l'} + \ldots + E_{n'}\,. \tag{25′}$$

is a solution with the same eigenvalue $E' = E$, when the sequence k', l', …, n' is a *permutation* of the numerical sequence k, l, …, n. If all of the numbers k, l, …, n are different, i.e., if (25) consists of nothing but different ψ-functions as factors, then the *one* eigenvalue:

$$E = E_k + E_l + \dots + E_n = E_{k'} + E_{l'} + \dots + E_{n'} = E_{k''} + E_{l''} + \dots + E_{n''} = \dots \tag{26}$$

will then be associated with a total of $N!$ mutually-distinct eigenfunctions:

$$\psi^{(1)} = \psi_k(\mathfrak{r}_1) \dots \psi_n(\mathfrak{r}_N), \qquad \psi^{(2)} = \psi_{k'}(\mathfrak{r}_1) \dots \psi_{n'}(\mathfrak{r}_N), \qquad \psi^{(N!)} = \dots, \tag{26$'$}$$

of the eigenvalue E, and one will refer to it as $N!$*-fold degenerate*. By contrast, if not all of the N numbers k, l, …, n are distinct, but they can be divided into groups of n_1, n_2, … mutually-equal numbers ($n_1 + n_2 + \dots = N$), then one will get only:

$$G = \frac{N!}{n_1!\, n_2! \cdots} \tag{26$''$}$$

distinct eigenfunctions $\psi^{(1)}$, $\psi^{(2)}$, …, $\psi^{(G)}$ for the *one* eigenvalue E under permutation of the sequence, so it will be G-fold degenerate. Finally, if all k, l, …, n are equal to each other then E will belong to only *one* eigenfunction $\psi_k(\mathfrak{r}_1)\, \psi_k(\mathfrak{r}_2) \dots \psi_n(\mathfrak{r}_N)$, so $E = E_k + E_l + \dots + E_n$ will be simply-degenerate and will usually just be called *nondegenerate*.

We now consider a G-fold eigenvalue. It belongs to not only the G eigenfunctions $\psi^{(1)}$, $\psi^{(2)}$, …, $\psi^{(G)}$, which arise from, say, $\psi^{(1)} = \psi_k(\mathfrak{r}_1)\, \psi_k(\mathfrak{r}_2) \dots \psi_n(\mathfrak{r}_N)$ by G permutations of the lower indices, but also to all linear combinations with arbitrary coefficients c :

$$\Psi = c_1 \psi^{(1)} + c_2 \psi^{(2)} + \dots + c_G \psi^{(G)}, \tag{27}$$

which are themselves once more eigenfunctions of that G-fold degenerate eigenvalue. Of course, they are linearly composed from the G eigenfunctions $\psi^{(1)}$ to $\psi^{(G)}$. However, one can define G eigenfunctions from the coefficients c_{ik} with a non-vanishing determinant:

$$\left.\begin{array}{l} \Psi^{(1)} = c_{11} \psi^{(1)} + c_{12} \psi^{(2)} + \dots + c_{1G} \psi^{(G)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \Psi^{(G)} = c_{G1} \psi^{(1)} + c_{G2} \psi^{(2)} + \dots + c_{GG} \psi^{(G)}, \end{array}\right\} \tag{28}$$

which are then linearly independent of each other. In that way, the c_{ik} in one of those rows can be chosen such that the function Ψ that arises will be *symmetric*, i.e., nothing will change under a permutation of one $\mathfrak{r}_K$ with another $\mathfrak{r}_L$. That will be the case when one chooses:

$$\Psi_{\text{sym.}} = \sum \psi_k(\mathfrak{r}_1)\psi_l(\mathfrak{r}_2)\cdots\psi_n(\mathfrak{r}_N)\,, \tag{29}$$

and sums over all G permutations of the sequence $k, l, \ldots, n$. $\Psi_{\text{sym.}}$ is then nothing but $\Psi = \psi^{(1)} + \psi^{(2)} + \cdots + \psi^{(G)}$ with the factors $c = 1$. In another row of (28), one can choose the c such that Ψ will become *antisymmetric*, i.e., it will always change sign when one switches any $\mathfrak{r}_K$ with any another $\mathfrak{r}_L$, namely:

$$\Psi_{\text{antisym.}} = \begin{vmatrix} \psi_k(\mathfrak{r}_1) & \psi_k(\mathfrak{r}_2) & \cdots & \psi_k(\mathfrak{r}_N) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \psi_n(\mathfrak{r}_1) & \psi_n(\mathfrak{r}_2) & \cdots & \psi_n(\mathfrak{r}_N) \end{vmatrix} \tag{30}$$

(with all c equal to + 1 or – 1), since switching $\mathfrak{r}_K$ with $\mathfrak{r}_L$ is equivalent to switching two columns in the determinant. One will then see that there can no longer be any symmetric or antisymmetric function among the remaining G – 2 composite function Ψ of (28), since otherwise not all G functions would be linearly independent. Moreover, $\Psi_{\text{antisym.}}$ will exist (and be non-zero) only when all functions ψ_k , ψ_l , …, ψ_n are *different*, because otherwise two or more rows in the determinant would be equal to each other, and that would make it vanish. $\Psi_{\text{antisym.}}$ will then be only $G = N!$-fold degenerate, where each of the N particles is in a different quantum state.

Whereas in the absence of mutual perturbations of the N particles, the coefficients c in (28) are arbitrary, things will be different when one introduces a perturbing potential that is symmetric in all particles and lets it converge to zero. One will indeed get the unperturbed solutions (28) again, but with well-defined coefficients c, and indeed, as perturbation theory shows, one will first get the symmetric solution, then the antisymmetric one, and finally G – 2 asymmetric solutions, which is a situation that **Heisenberg** recognized to be an analogue of the *resonance phenomena* in the classical mechanics of coupled systems. In particular, in the two-body problem, only the symmetric and antisymmetric solutions will remain:

$$\left.\begin{aligned} \Psi_{\text{sym.}} &= \frac{1}{\sqrt{2}}[\psi_k(\mathfrak{r}_1)\psi_l(\mathfrak{r}_2) + \psi_l(\mathfrak{r}_1)\psi_k(\mathfrak{r}_1)], \quad E = E_k + E_l, \\ \Psi_{\text{antisym.}} &= \frac{1}{\sqrt{2}}[\psi_k(\mathfrak{r}_1)\psi_l(\mathfrak{r}_2) - \psi_l(\mathfrak{r}_1)\psi_k(\mathfrak{r}_1)], \quad E = E_k + E_l \end{aligned}\right\} \tag{31}$$

for an asymptotically-vanishing perturbing potential. The factor $1/\sqrt{2}$ is included in order to normalize to 1. For a different eigenvalue E', one will have:

$$\left.\begin{aligned} \Psi_{\text{sym.}} &= \frac{1}{\sqrt{2}}[\psi_{k'}(\mathfrak{r}_1)\psi_{l'}(\mathfrak{r}_2)+\psi_{l'}(\mathfrak{r}_1)\psi_{k'}(\mathfrak{r}_1)], \quad E' = E_{k'} + E_{l'}, \\ \Psi_{\text{antisym.}} &= \frac{1}{\sqrt{2}}[\psi_{k'}(\mathfrak{r}_1)\psi_{l'}(\mathfrak{r}_2)-\psi_{l'}(\mathfrak{r}_1)\psi_{k'}(\mathfrak{r}_1)], \quad E' = E_{k'} + E_{l'}. \end{aligned}\right\} \tag{31′}$$

If the perturbing potential that is symmetric in the particles increases then the symmetry (antisymmetry, resp.) will be preserved, but now the expressions for the right-hand sides of (31), (31′) will be modified.

32. **Forbidden combinations of systems of terms**. – We shall now prove the theorem, which is important in the special case of a two-electron system, that a transition from a state that is described by a symmetric eigenfunction to an antisymmetric state is forbidden. In order to do that, we will form the stated equation:

$$\left.\begin{aligned} &\int \Psi_{\text{sym.}}(\mathfrak{r}_1,\mathfrak{r})\cdot\Psi'_{\text{antisym.}}(\mathfrak{r}_1,\mathfrak{r})\,dv_1 + \int \Psi_{\text{sym.}}(\mathfrak{r},\mathfrak{r}_2)\cdot\Psi'_{\text{antisym.}}(\mathfrak{r},\mathfrak{r}_2)\,dv_2 \\ = \;& -\int \Psi_{\text{sym.}}(\mathfrak{r},\mathfrak{r}_1)\cdot\Psi'_{\text{antisym.}}(\mathfrak{r},\mathfrak{r}_1)\,dv_1 - \int \Psi_{\text{sym.}}(\mathfrak{r}_2,\mathfrak{r})\cdot\Psi'_{\text{antisym.}}(\mathfrak{r}_2,\mathfrak{r})\,dv_2 \end{aligned}\right\} \tag{32}$$

with the help of (31), (31′). The left-hand side of that is the amplitude of the *transition density* (21″), which is definitive of the transition from $\Psi_{\text{sym.}}$ to $\Psi'_{\text{antisym.}}$.

If one now formally switches the indices 1 and 2 on the right-hand side then that will show that the right-hand side is also simultaneously the *negative* of the left-hand side. Both sides must be equal to zero then. That proves the vanishing of the transition density, from which the fact that the combination is forbidden will follow (conclusion of no. **30**). If the electron system is, say, in a state with a symmetric eigenfunction at some point in time then it will remain in it for all time when the transition to an antisymmetric or asymmetric state is forbidden. Now, experiments show that in reality only the states with *antisymmetric eigenfunctions* (cf., *infra*) will occur, which can also never go to symmetric or asymmetric states, due to the fact that the combination is forbidden.

If one, in fact, allows only antisymmetric eigenfunctions then one will restrict oneself to those states of the total system in which (in the absence of coupling) every particle will lie in a different quantum state, because otherwise the determinant (30) would vanish. Now, **Pauli** proposed the principle that every electron in a multi-electron system must be (in the absence of coupling) intrinsically in a different quantum state. Here, the **Pauli** principle will then prove to be equivalent to the demand that only the *antisymmetric* eigenfunctions are allowed.

The question of how to decompose multi-electron systems into non-combining systems of terms was solved conclusively by **E. Wigner** and **F. Hund** (1) once **Heisenberg** (2) and **Dirac** (3) had posed the basic questions in the simplest cases and clarified them.

As an example, following **Heisenberg**, we shall introduce the system of terms for *helium*, which consists of a positive nucleus and two electrons whose **Goudsmit-Uhlenbeck** spin impulses have two different possible directions, and can therefore give rise to two different states, which we would like to distinguish by the signs + and –. In the absence of mutual perturbations of the two electrons, the solution ψ to the **Schrödinger** equation will split into a product $\varphi \cdot \chi$, in which φ and χ mean two solutions of the one-electron problem, and indeed according to the direction of the spin axis, we shall distinguish between φ^+ and φ^-, and χ^+ and χ^-. Upon considering the mutual perturbation of the electrons, the following eight solutions ψ to the two-electron problem will then remain:

$$\left.\begin{array}{ll}\varphi^+(\mathfrak{r}_1)\cdot\chi^+(\mathfrak{r}_2)-\chi^+(\mathfrak{r}_1)\cdot\varphi^+(\mathfrak{r}_2), & \varphi^-(\mathfrak{r}_1)\cdot\chi^-(\mathfrak{r}_2)-\chi^-(\mathfrak{r}_1)\cdot\varphi^-(\mathfrak{r}_2),\\ \varphi^+(\mathfrak{r}_1)\cdot\chi^-(\mathfrak{r}_2)-\chi^-(\mathfrak{r}_1)\cdot\varphi^+(\mathfrak{r}_2), & \varphi^-(\mathfrak{r}_1)\cdot\chi^+(\mathfrak{r}_2)-\chi^+(\mathfrak{r}_1)\cdot\varphi^-(\mathfrak{r}_2),\end{array}\right\}\quad\text{antisymmetric}$$

$$\left.\begin{array}{ll}\varphi^+(\mathfrak{r}_1)\cdot\chi^+(\mathfrak{r}_2)+\chi^+(\mathfrak{r}_1)\cdot\varphi^+(\mathfrak{r}_2), & \varphi^-(\mathfrak{r}_1)\cdot\chi^-(\mathfrak{r}_2)+\chi^-(\mathfrak{r}_1)\cdot\varphi^-(\mathfrak{r}_2),\\ \varphi^+(\mathfrak{r}_1)\cdot\chi^-(\mathfrak{r}_2)+\chi^-(\mathfrak{r}_1)\cdot\varphi^+(\mathfrak{r}_2), & \varphi^-(\mathfrak{r}_1)\cdot\chi^+(\mathfrak{r}_2)+\chi^+(\mathfrak{r}_1)\cdot\varphi^-(\mathfrak{r}_2).\end{array}\right\}\quad\text{symmetric}$$

The four symmetric solutions among them are excluded by the **Pauli** principle. Of the four antisymmetric solutions, the two in the first row are possessed by a state of He in which *both* electrons have *parallel* spin impulses (+ + or – –, resp.), while the second row belongs to *antiparallel* spin impulses (+ – or – +, resp.). The first row corresponds to a *triplet term*, while the one corresponds to a *singlet term* (both of them represent two ψ-functions that belong to the same eigenvalue). In particular, if both electrons "move on equivalent orbits" (i.e., when $\varphi = \chi$) then the two ψ-functions in the first row will be equal to zero, and the function in the second row:

$$\varphi^+(\mathfrak{r}_1)\varphi^-(\mathfrak{r}_2)-\varphi^-(\mathfrak{r}_1)\varphi^+(\mathfrak{r}_2)\ , \qquad \varphi^-(\mathfrak{r}_1)\varphi^+(\mathfrak{r}_2)-\varphi^+(\mathfrak{r}_1)\varphi^-(\mathfrak{r}_2)$$

will be non-zero. Equivalent orbits will then give only singlet terms. For example, only singlet terms will appear in the *ground state* of He (viz., both electrons are on orbits with the azimuthal quantum number $l = 0$), while the higher excited terms ($l_1 = 0$, $l_2 = 1, 2, \ldots$) define singlet terms (para-helium) and triplet terms (ortho-helium).

---

(1) **E. Wigner**, Zeit. Phys. **40** (1926), pps. 492 and 883; *ibid.*, **43** (1927), pp. 524. – **F. Hund**, Zeit. Phys. **43** (1927), pp. 788.

(2) **W. Heisenberg**, Zeit. Phys. **38** (1926), pp. 411; *ibid.*, **39** (1926), pp. 499. – A related resonance process was also used as the basis for homopolar molecular couplings by **Heitler** and **London**, Zeit. Phys. **44** (2917), pp. 455.

(3) **P. Dirac**, Proc. Roy. Soc. **112** (1926), pp. 661.

33. Bose and Fermi statistics. – In order to derive the thermodynamical properties of a system that is composed of N equivalent particles, one must first establish the *statistical weight* of the individual "states" of the system.

N mutually-noninteracting particles (e.g., ideal gas atoms or electrons around a nucleus with no coupling forces) might be distributed over the total energy E in the following way:

$$E = n_1 E_1 + n_2 E_2 + \dots , \qquad n_1 + n_2 + \dots = N ,$$

in which E_k shall be the eigenvalue that belongs to the unperturbed eigenfunction ψ_k . According to (26″), the state of the total system that is characterized by the numbers (n_1, n_2, …) can then be represented in:

$$G = \frac{N!}{n_1!\, n_1! \cdots}$$

ways, namely, by G linearly-independent functions $\psi^{(1)}$, $\psi^{(2)}$, …, $\psi^{(G)}$. If one regards all of those G states as *distinct* and counts each of them with the statistical weight of 1 then the state that is characterized by the numbers (n_1, n_2, …) will possess the total weight G. **Boltzmann** statistics also ascribes the same weight G to a state (n_1, n_2, …) in which each of the G distributions of N particles over the energy levels E_1, E_2, … is counted as an equally-probable state.

According to **Bose**, in order to obtain **Planck**'s law of radiation, in the distribution of $M = n_1 + n_2 + \dots$ light quanta over the energy levels E_1, E_2, … one must count the distribution state that is characterized by the numbers (n_1, n_2, …) with weight 1. The G **Boltzmann** permutations of the individual light particles will then be regarded as indistinguishable in that way. If one, with **Einstein**, adapts **Bose** statistics to material particles (e.g., gas atoms, atomic electrons) then that will mean that of the G linearly-independent state functions $\psi^{(1)}$, $\psi^{(2)}$, …, $\psi^{(G)}$ of the particle system, only one of then will be allowable, and it will be counted with the statistical weight 1. One will then remain consistent with **Bose-Einstein** statistics when one allows only the *symmetric* state function $\psi_{\text{sym.}}$ to be that one. Indeed, it will not change under permutations of the individual particles [i.e., switching $\mathfrak{r}_K$ with $\mathfrak{r}_L$ in (29)]. Now, **Bose** statistics (the equivalent choice of *symmetric* eigenfunctions in the undulatory picture) are indeed exhibited by light, but not by matter.

The **Pauli** principle that no quantum state can possess more than one electron (†) is true for material particles (at least for the N-electron system of an atom). In the undulatory picture, that means that the unperturbed eigenfunctions ψ_k of the N electrons must all vanish ($E = E_1 + E_2 + \dots + E_N$, $N = 1 + 1 + \dots + 1$, $G = N!$), or that the sequence (27) must *vanish*, resp., whenever not all ψ_k are different. However, that will be the case only when one allows the *antisymmetric* eigenfunction (30).

Fermi has adapted the **Pauli** principle to the N particles of an ideal gas, of which one then demands that it must be found in N different energy cells. In the undulatory picture, that means that of the $G = N!$ state functions $\Psi^{(1)}$, …, $\Psi^{(G)}$ of the total system, *only one* of them can be

---

(†) Translator: I believe he means "more one electron in the same spin state."

regarded as existing, and indeed that will be the antisymmetric state function (30), because in that way, distributions of the N particles over less than N energy cells (i.e., some energy cells being occupied by more than one particle) will be excluded [because the determinant (30) will vanish when two or more cells are equal]. The **Pauli** principle and the **Fermi** statistics of material particles are then characterized by the demand that only antisymmetric state functions are allowed and must be counted with a statistical weight of 1.

34. Connection with matrix mechanics (1). – Here, a somewhat-different representation of quantum theory will be touched upon that will first be discussed in connection with Chapter VII, namely, the **Heisenberg**, **Born**, and **Jordan** approach to quantum mechanics.

We understand f to mean an *operator*, which we define by:

$$\{f, \psi_m\} = \sum_n f^{nm} \cdot \psi_n , \tag{33}$$

i.e., when the operation f is performed on a function ψ_m , the result shall be a series development in terms of functions ψ_n with certain constant coefficients (viz. generalized ***Fourier** coefficients*) that will first establish the sense of operation f when one gives precise values to those coefficients. If the functions ψ_n (q) are orthogonal to each other and normalized to 1 then we can inversely determine the coefficients f^{mn} from (33), which we shall also refer to as *matrix elements* of f, to be:

$$f^{mn} = \int \psi_n \{f, \psi_m\}\, dv , \tag{34}$$

when integrated over the entire space of the coordinates q_K . The Schrödinger equation:

$$\{H, \psi_m\} = E_m \psi_m ,$$

with the operator H, can also be written in the form:

$$\sum_n H^{nm} \psi_n = E_m \psi_m ,$$

from which it will emerge that the matrix elements of H will be equal to:

$$H^{nm} = 0 \qquad \text{for} \qquad n \neq m , \qquad\qquad H^{mm} = E_m . \tag{35}$$

The Hamiltonian function H is a function of the p_K and q_K, which are coupled by finite or infinitely-many additions and multiplications in it. In order to compose the matrix elements H^{nm} from the

---

(1) **E. Schrödinger**, Ann. Phys. (Leipzig) **79** (1926), pp. 734, and independently of him, **C. Eckart**, Phys. Rev. **28** (1926), pp. 711.

matrix elements of the p_K and q_K, one must then know how the matrix elements of two operators f and g show up in the formation of the matrix elements of the operators $f + g$ and $f \cdot g$. Now, from (34), one has:

$$f^{nm} = \int \psi_n \, f \, \psi_m \, dv , \quad g^{nm} = \int \psi_n \, g \, \psi_m \, dv , \quad (f\,g)^{nm} = \int \psi_n \, f \, g \, \psi_m \, dv .$$

On the other hand, from a theorem in the theory of eigenfunctions (^{1}):

$$\int \psi_n \, f \, g \, \psi_m \, dv = \sum_l \int \psi_n \, f \, \psi_l \, dv \cdot \int \psi_l \, g \, \psi_m \, dv ,$$

which will then give:

$$(f\,g)^{nm} = \sum_l f^{nl} \, g^{lm} \tag{36}$$

as the *product rule*. That is equivalent to the rule that the elements of a product of two *determinants* are composed from the elements of the original determinants. The *sum rule* reads simply:

$$(f \pm g)^{nm} = \int \psi_n \, (f \pm g) \psi_m \, dv = \int \psi_n \, f \, \psi_m \, dv \pm \int \psi_n \, g \, \psi_m \, dv = f^{nm} \pm g^{nm} . \tag{36′}$$

Now, one can then construct the *matrix elements* F^{ik} of a function $F(p, q)$ that is formed from the p and q by summation and products of the *matrix elements* p_K^{mn} and q_K^{mn}.

For **Schrödinger**, the operator p_K shall mean $\frac{h}{2i\pi} \frac{\partial}{\partial q_K}$. Corresponding to (33), one can then infer here that:

$$\frac{h}{2i\pi} \frac{\partial}{\partial q_K} \psi_m = \{p_K, \psi_m\} = \sum_n p_K^{nm} \, \psi_n ,$$

and conversely, corresponding to (34), the values of its matrix elements will be:

$$p_K^{nm} = \int \psi_n \{p_K \, \psi_m\} dv = \frac{h}{2i\pi} \int \psi_n \frac{\partial}{\partial q_K} \psi_m \, dv .$$

We further consider the *operator* $p_K \, q_K - q_K \, p_K$. One has:

$$\frac{h}{2i\pi} \left\{ \frac{\partial}{\partial q_K} q_K \, \psi_m - q_K \frac{\partial}{\partial q_K} \psi_m \right\} = \{p_K \, q_K - q_K \, p_K, \psi_m\} = \sum_n (p_K \, q_K - q_K \, p_K)^{nm} \, \psi_n ,$$

and conversely:

(^{1}) That "completeness relation" will not be employed directly in the derivation of the same result in no. **65**.

$$(p_K\, q_K - q_K\, p_K)^{nm} = \int \psi_n \cdot \frac{h}{2i\pi}\left\{\frac{\partial}{\partial q_K} q_K\,\psi_m - q_K \frac{\partial}{\partial q_K}\psi_m\right\} dq \;=\; \frac{h}{2i\pi}\int \psi_n \cdot \psi_m\, dv$$

$$= \frac{h}{2i\pi}\cdot 1 \qquad \text{for} \qquad n = m \quad \text{and} \qquad = 0 \qquad \text{for } n \neq m.$$

The series development with just one term then follows from that:

$$\{p_K\, q_K - q_K\, p_K,\ \psi_m\} = \frac{h}{2i\pi}\cdot \psi_m\,, \tag{37}$$

i.e., the action of the operator $p_K\, q_K - q_K\, p_K$ on ψ_m is equivalent to *multiplication by the factor* $h / 2i\pi$. One easily shows, correspondingly:

$$\left.\begin{array}{lll} (p_K\, q_L - q_L\, p_K)^{nm} = 0 & \text{for } K \neq L, \text{so} & \{p_K\, q_L - q_L\, p_K, \psi_m\} = 0 \cdot \psi_m\,, \\ (p_K\, p_L - p_L\, p_K)^{nm} = 0 & \text{for all } K \text{ and } L, \text{so} & \{p_K\, p_L - p_L\, p_K, \psi_m\} = 0 \cdot \psi_m\,, \\ (q_K\, q_L - q_L q_K)^{nm} = 0 & \text{for all } K \text{ and } L, \text{so} & \{q_K\, q_L - q_L q_K, \psi_m\} = 0 \cdot \psi_m\,, \end{array}\right\} \tag{37′}$$

i.e., the latter operators are equivalent to *multiplication by the factor zero*.

Summarizing (37), (37′), we then find that the operator $p_K\, p_L$ is indeed equal to the operator $p_L\; p_K$, and $q_K\, q_L = q_L\, q_K$, and finally $p_K\, q_L = q_L\, p_K$ (for $K \neq L$), while $p_K\, q_K$ is not equal to $q_K\, p_K$, since $p_K\, q_K - q_K\, p_K$ is not, in fact, equal to zero, but to the operator h / $2i\ \pi$ as a product factor. That can be expressed somewhat differently by saying that whereas the operator p_L *commutes* with the operator p_K, q_L commutes with q_K, and p_L commutes with q_K (for $K \neq L$), p_K *does not commute* with q_K, and indeed we have the *commutation relations:*

$$\left.\begin{array}{l} p_K\, q_L - q_L\; p_K = \begin{cases} \dfrac{h}{2i\pi} & \text{for} \quad K = L, \\ 0 & \text{for} \quad K \neq L, \end{cases} \\ p_K\; p_L - p_L\; p_K = 0, \quad q_K\, q_L - q_L\, q_K = 0. \end{array}\right\} \tag{38}$$

The **Schrödinger** eigenvalue problem is then equivalent to the following **Heisenberg-Born-Jordan** problem: Construct the matrix elements H^{mn} of the operator H (p, q) from the matrix elements p_K^{mn} and q_K^{mn}. If one observes the commutation rules (37), (37′):

$$\left.\begin{array}{l}\sum_l p_K^{nl}\,q_L^{lm} - q_L^{nl}\,p_K^{lm} = \begin{cases}\dfrac{h}{2i\pi} & \text{for} \quad n = m \\ 0 & \text{for} \quad n \neq m\end{cases} \text{for } K = L, = 0 \text{ for } K \neq L, \\ \sum_l p_K^{nl}\,p_L^{lm} - p_L^{nl}\,p_K^{lm} = 0, \quad \sum_l q_K^{nl}\,q_L^{lm} - q_L^{nl}\,q_K^{lm} = 0\end{array}\right\} \tag{39}$$

then one seeks to assign constant values to the p_K^{mn} and q_K^{mn} such that:

$$H^{mn} = 0 \qquad \text{for} \qquad m \neq n\,,$$

as in (35), and only the H^{mm} will take on certain non-zero values. Those will then be the desired eigenvalues E_m of the energy for the quantum-mechanical problem.

CHAPTER IV

UNDULATORY MECHANICS OF TIME-VARYING SYSTEMS

35. Time-varying potential. – Whereas the *stationary* states of a system of mass-points in a potential field $U(q)$ that did not depend upon time were treated in Chapter III, the *transition* between stationary states shall also brought under consideration now when they are created by time-dependent fields. If one now puts, e.g., the interaction energy between a system of electrons and incident light into a form in which one employs the conservative potential $U(q)$ of the electron system, extended by the time-dependent potential $V(q, t)$ of the light field, in a classical way in order to exhibit the **Hamiltonian** function and reinterprets the fundamental classical equation in terms of the wave-theoretic one then, to the extent that one forgets the reaction of the electron system on the radiation, it will be inconsequential whether the classical theory takes the radiation damping into account, but in that way it is not at all representable by a potential function that depends upon position and time, but rather it will depend upon the state of the motion of the electron system. We will first see how to overcome that absence in Chapter V, where the interaction of the light with the atom will be treated as a legitimate part of the total system of atom + light. Before we do that, we shall consider the simpler case, which is sufficient for many purposes, in which the total energy of the mass-point system is described classically by a kinetic energy and a potential one that shall depend upon time, in addition to position (e.g., a light field with no consideration given to the reaction).

It is known from classical mechanics that the time coordinate t is associated with negative energy $(-E)$ as a harmonically-conjugate "impulse." Whereas the conservative energy equation:

$$H(q, p) - E = 0 , \tag{1}$$

when one appeals to an action function $S(q)$ and:

$$p_1 = \frac{\partial S}{\partial q_1}, \qquad \ldots, \qquad p_N = \frac{\partial S}{\partial q_N} \qquad (p_K \text{ is conjugate to } q_K) \tag{1$'$}$$

will go to the conservative **Hamilton-Jacobi** differential equation for S :

$$H\left(q, \frac{\partial S}{\partial q}\right) - E = 0 , \tag{1$''$}$$

the non-conservative energy equation:

$$H(q, t, p) - E = 0 , \tag{2}$$

with the help of an action function $S(q, t) = S(q) - E\,t$ and:

$$p_1 = \frac{\partial S}{\partial q_1}, \qquad \ldots, \qquad p_N = \frac{\partial S}{\partial q_N}, \qquad -E = \frac{\partial S}{\partial t}, \tag{2′}$$

will go to the non-conservative **Hamilton-Jacobi** differential equation for S :

$$H\left(q, t, \frac{\partial S}{\partial q}\right) + \frac{\partial S}{\partial t} = 0\,. \tag{2″}$$

Now, the introduction of *wave mechanics* consists of saying that one replaces the quantities (2′) in the **Hamilton-Jacobi** equation with the *operators:*

$$p_K \sim \frac{h}{2i\pi}\frac{\partial}{\partial q_K}, \qquad -E \sim \frac{h}{2i\pi}\frac{\partial}{\partial t}, \tag{3}$$

and applies them to a coordinate function $\tilde{\psi}$ (q_1, …, q_N, t):

$$\left\{H\left(q, t, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + \frac{h}{2i\pi}\frac{\partial}{\partial t}, \tilde{\psi}\right\} = 0\,. \tag{4}$$

If one composes the classical Hamiltonian function H from the potential energy U (q, t) and kinetic energy, and the latter is a quadratic form in the impulse (non-relativistic mechanics), as in no. **22**, then one will have, from (4):

$$\left(\frac{h}{2i\pi}\right)^2 \frac{1}{2\mu}\Delta\tilde{\psi} + U\tilde{\psi} + \frac{h}{2i\pi}\frac{\partial\tilde{\psi}}{\partial t} = 0\,,$$

or, when written differently:

$$\Delta\tilde{\psi} - \frac{8\pi^2\mu}{h^2}U\,\tilde{\psi} + \frac{4\pi i\mu}{h}\frac{\partial\tilde{\psi}}{\partial t} = 0 \qquad \text{(wave equation)}, \tag{5}$$

for the determination of $\tilde{\psi}(q, t)$.

If $\tilde{\psi}$ is a solution of (5) then the complex-conjugate function ψ will be a solution of the conjugate equation:

$$\boxed{\Delta\psi - \frac{8\pi^2\mu}{h^2}U\,\psi - \frac{4\pi i\mu}{h}\frac{\partial\psi}{\partial t} = 0 \quad \text{(wave equation).}} \tag{5′}$$

One can also derive the wave equations (5), (5′) from a *variational principle* (cf., no. **25**): In the "**Lagrangian** function" of the classical mechanical problem:

$$L = \text{kin.} - \text{pot.} + \frac{\partial S}{\partial t} = T\left(q, \frac{\partial S}{\partial q}\right) - U(q,t) \pm \frac{\partial S}{\partial t}\,, \tag{6}$$

one uses the Ansatz $\psi = e^{2i\pi S/h}$ to replace:

$$\left.\begin{array}{l} \dfrac{\partial S}{\partial q_K} \quad \text{with} \quad \dfrac{h}{2i\pi}\dfrac{1}{\psi}\dfrac{\partial \psi}{\partial q_K} \quad \text{or} \quad -\dfrac{h}{2i\pi}\dfrac{1}{\tilde{\psi}}\dfrac{\partial \tilde{\psi}}{\partial q_K}, \text{resp.,} \\ \text{and} \\ +\dfrac{\partial S}{\partial t} \quad \text{with} \quad \dfrac{1}{2}\left(\dfrac{h}{2i\pi}\dfrac{1}{\psi}\dfrac{\partial \psi}{\partial q_K} - \dfrac{h}{2i\pi}\dfrac{1}{\tilde{\psi}}\dfrac{\partial \tilde{\psi}}{\partial q_K}\right), \end{array}\right\} \tag{6′}$$

and one then solves the variational problem:

$$J = \iint \psi\tilde{\psi} \cdot L\, dt \cdot dv = \text{extremum}, \tag{6″}$$

i.e., when written out in detail:

$$\delta J = \delta \iint \left[\sum_K \sum_L \frac{h^2}{8\pi^2 \mu} g_{KL} \frac{\partial \psi}{\partial q_K}\frac{\partial \tilde{\psi}}{\partial q_L} + U\,\psi\tilde{\psi} + \frac{h}{4i\pi}(\tilde{\psi}\dot{\psi} - \psi\dot{\tilde{\psi}})\right]\sqrt{g^*}\, dq_1 \cdots dq_N\, dt = 0\,. \tag{6‴}$$

If one denotes the integrand, including the factor $\sqrt{g^*}$, by F here then the two **Euler** equations:

$$\sum_L \frac{d}{dq_L}\left(\frac{\partial F}{\partial \tilde{\psi}_{q_L}}\right) + \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{\tilde{\psi}}}\right) - \frac{\partial F}{\partial \tilde{\psi}} = 0\,,$$

$$\sum_K \frac{d}{dq_K}\left(\frac{\partial F}{\partial \psi_{q_K}}\right) + \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{\psi}}\right) - \frac{\partial F}{\partial \psi} = 0$$

will be identical to the wave equations (5), (5′).

Now let *U not depend upon t*, in particular, i.e., $U = U_0(q)$. If one then introduces the Ansatz $\psi(q, t) = \psi(q) \cdot e^{-2i\pi Et/h}$ [$\tilde{\psi}(q,t) = \psi(q) \cdot e^{+2i\pi Et/h}$, resp.] into (5) [(5′), resp.] then what will remain for $\psi(q)$ is the previous oscillation equation [Sec. III, (9)]:

$$\Delta\psi - \frac{8\pi^2 \mu}{h^2}(U - E)\,\psi = 0 \qquad \text{(oscillation equation)} \tag{7}$$

with the solutions ψ_n for the eigenvalue E_n such that the desired solution of the wave equation will read:

$$\psi_n(q, t) = \psi_n(q) \cdot e^{-2i\pi E_n t/h}\,. \tag{8}$$

[Observe that when $\psi_n(q)$ is a solution to the real oscillation, $\tilde{\psi}_n(q)$ will also be a solution to the same equation with the same eigenvalue.]

However, in the case of $U = U_0(q)$, from the form of the wave equation (5) [(5′), resp.], it will possess not only the solution (8) with the index n, but also all linear combinations of such solutions with differing n and arbitrary constant coefficients a_n :

$$\psi(q, t) = \sum a_n \psi_n(q) \cdot e^{-2i\pi E_n t/h} . \tag{9}$$

However, if *U also depends upon t*, so $U = U_0(q) + U_1(q, t)$, with the time-varying perturbing potential U_1 then one can always exhibit $\psi(q, t)$ in the form of the series (9) as a development in the eigenfunctions $\psi_n(q)$, but with the time-dependent coefficients $a_n(t)$:

$$\psi(q, t) = \sum a_n(t) \psi_n(q) \cdot e^{-2i\pi E_n t/h} , \quad \tilde{\psi} = \sum a_n(t) \psi_n(q) \cdot e^{2i\pi E_n t/h} . \tag{10}$$

The physical meaning of such a composite solution, in particular the meaning of the coefficients a_n, with and without time dependency, will be explained in the following no. **36**.

36. Conservation laws. – If one multiplies (5) by $\tilde{\psi}$ and (5) by ψ and forms, in one case, the sum and in the other, the difference of the equations thus-obtained then one will get the two equations:

$$\operatorname{div}(\tilde{\psi} \operatorname{grad} \psi - \psi \operatorname{grad} \tilde{\psi}) - \frac{4\pi i \mu}{h} \frac{\partial}{\partial t} |\psi|^2 = 0 \tag{11}$$

and

$$\left.\begin{aligned} &\operatorname{div}(\tilde{\psi} \operatorname{grad} \psi - \psi \operatorname{grad} \tilde{\psi}) - \frac{4\pi i \mu}{h} \frac{\partial}{\partial t} |\psi|^2 \\ &\quad = 2\left\{ |\operatorname{grad} \psi|^2 + \frac{8\pi^2 \mu}{h^2} U \cdot |\psi|^2 \right\}, \end{aligned}\right\} \tag{12}$$

when one uses the following formulas:

$$\Delta\varphi = \operatorname{div} \operatorname{grad} \varphi \qquad \text{and} \qquad \chi\,\Delta\varphi = \operatorname{div}(\chi \operatorname{grad} \varphi) - \operatorname{grad} \chi \operatorname{grad} \varphi , \tag{13}$$

which are also true in multi-dimensional space with a non-Euclidian metric. If one now integrates (11) and (12) with $dv = \sqrt{g^*}\, dq_1\, dq_2 \ldots$ and considers suitable boundary conditions for ψ at infinity (periodicity conditions on the angle coordinates, resp.) then the contribution $\int \operatorname{div}(\cdots)\, dv$ will drop out with the use of **Gauss**'s theorem, and what will remain are the following equations [viz., the **Schrödinger** and **Born** *conservation laws* (¹)]:

(¹) **E. Schrödinger**, Ann. Phys. (Leipzig) **81** (1926), pp. 109; **M. Born**, Zeit. Phys. **40** (1927), pp. 167.

$$\frac{d}{dt}\int |\psi|^2 \, dv = 0 \,, \tag{14}$$

$$\int \left(\frac{h^2}{8\pi^2 \mu} |\operatorname{grad} \psi|^2 + U(q,t) |\psi|^2 \right) dv = \frac{h}{4i\pi} \int \left(\psi \frac{\partial \tilde{\psi}}{\partial t} - \tilde{\psi} \frac{\partial \psi}{\partial t} \right) dv \,. \tag{15}$$

Next, (14) gives us the right to normalize:

$$\int \psi \tilde{\psi} \, dv = 1 \,, \tag{14′}$$

since the value 1 does not, in fact, change in time. From now on, we can interpret $\psi\tilde{\psi} = |\psi|^2$ as the *relative density*, $\varepsilon |\psi|^2 = \rho_\varepsilon$ as the charge density, $\mu |\psi|^2 = \rho_\mu$ as the mass density in coordinate space in the state that is represented by $\psi(q, t)$. (14) will then imply that the total amount of charge and mass will be conserved, even though the density at the individual locations q in coordinate space changes in time.

If one now uses the series development (10) for ψ then one will have:

$$\psi\tilde{\psi} = \sum |a_n|^2 |\psi_n|^2 + \sum\sum a_n \tilde{a}_m \psi_n \tilde{\psi}_m \cdot e^{-2i\pi(E_n - E_m)t/h} \,. \tag{16}$$

If one integrates that, while recalling (14′) and the orthogonality and normality of the ψ_n, then one will have:

$$1 = \int \psi \tilde{\psi} \, dv = |a_1|^2 + |a_2|^2 + |a_3|^2 + \dots \tag{16′}$$

One can correspondingly regard ψ as a representative of a state in which either each individual atom is just as well in the quantum state 1 as 2, etc., or simultaneously in all quantum states, but distributed over the individual states with the relative weights $|a_1|^2$, $|a_2|^2$, …, or also a state in which many of the atoms are in state 1, while the others are in state 2, etc., and with the relative frequencies a_1^2, a_2^2, etc. With the first picture, according to (10), ψ represents a *beat* (*Schwebung*) of the oscillations that belong to the individual states with the individual amplitudes $a_n(t)$. The total density $|\psi|^2$ that prevails at each location q in space is therefore not equal to the sum $\sum a_n^2 |\psi_n|^2$ of the partial differences $|\psi_n|^2$ of the individual quantum states, provided with the weighting factors $|a_n|^2$, but as (16) shows, it is further augmented by a double sum that is ascribed to a typical *interference* of the partial oscillations: One does not add the partial densities together, but one *superposes* the partial oscillations $a_n \psi_n$ into the total oscillation amplitude ψ, whose square first gives the total density afterwards.

If one substitutes the series (10) in the right-hand side of (15) then it will reduce to:

$$\sum E_n |a_n|^2 \,.$$

We would like to regard that as the total energy E of the state that ψ represents, which will give us all the more right to also interpret the left-hand side of (15):

$$\int U(q,t)|\psi|^2\,dv = U$$

as the total *potential energy* in the state ψ.

37. Hydrodynamical interpretation. – Along with the spatially-distributed potential energy, yet another term appears on the left-hand side of (15) that can be interpreted as a spatially-distributed *kinetic* energy of a current. However, as **Madelung** showed (1), only part of the term can be regarded as a kinetic energy of a current. The remainder can be interpreted as an "internal" stress energy (cf., *infra*). The current density j that belongs to the *density* $\rho = |\psi|^2$ can be read off from (11) when one regards (11) as a *hydrodynamical equation of continuity:*

$$\operatorname{div} j + \frac{\partial\rho}{\partial t} = 0 \qquad \text{with} \qquad \left.\begin{aligned} &\rho = \psi\tilde{\psi} = |\psi|^2, \\ &j = \frac{h}{4i\pi\mu}(\psi\,\operatorname{grad}\tilde{\psi} - \tilde{\psi}\,\operatorname{grad}\psi). \end{aligned}\right\} \tag{17}$$

One obtains the mass and charge density from ρ upon multiplying by μ and ε.

Madelung arrived at an especially clear form for the hydrodynamical representation when he put the solution ψ to the wave equation (6):

$$\Delta\psi - \frac{8\pi^2\mu}{h^2}U\psi + i\frac{4\pi\mu}{h}\frac{\partial\psi}{\partial t} = 0$$

into the form that has been customary since the time of **de Broglie**'s mechanics:

$$\psi = \alpha\cdot e^{2i\pi S/h}, \tag{18}$$

in which the *real* quantities α and S both depend upon q and t. Conversely, the hydrodynamical meaning of α and S will then follow from (17), (18):

$$\alpha = |\psi| = \sqrt{\rho}, \qquad S = \frac{h}{4i\pi}\ln\left(\frac{\psi}{\tilde{\psi}}\right) = \text{real part of } \frac{h}{4i\pi}\ln\psi. \tag{18′}$$

Introducing the Ansatz (18) into the wave equation and separating the real and imaginary components will lead to the following two equations for α and S :

---

(1) **E. Madelung**, "Quantentheorie in hydrodynamischer Form," Zeit. Phys. **40** (1926), pp. 322.

$$\frac{h^2}{4\pi^2}\,\Delta\alpha - \alpha\cdot(\text{grad}\,S)^2 - 2\mu\alpha U + 2\mu\alpha\,\frac{\partial S}{\partial t} = 0 \tag{18″}$$

and

$$\alpha\,\Delta S + 2\,(\text{grad}\,\alpha\cdot\text{grad}\,S) - 2\mu\frac{\partial\alpha}{\partial t} = 0\,. \tag{18‴}$$

After multiplying by a, one can also write the last equation in the form:

$$\text{div}\left(\frac{\alpha^2}{\mu}\,\text{grad}\,S\right) + \frac{\partial\alpha^2}{\partial t} = 0\,.$$

It will then represent a hydrodynamical continuity equation [cf., (17)] with:

$$\left.\begin{aligned} &\rho = \alpha^2 = \text{density}, \quad j = \frac{\alpha^2}{\mu}\,\text{grad}\,S = \text{current density}, \\ &\mathfrak{v} = \frac{j}{\rho} = \frac{1}{\mu}\,\text{grad}\,S = \text{flow velocity}. \end{aligned}\right\} \tag{19}$$

On the other hand, if one divides (18″) by $2\alpha\,\mu^2$ and takes its gradient then one will get the relation:

$$-\frac{\text{grad}\,U}{\mu} + \frac{h^2}{8\pi^2\mu^2}\,\text{grad}\,\frac{\Delta\alpha}{\alpha} = \frac{\partial\mathfrak{v}}{\partial t} + \tfrac{1}{2}\,\text{grad}\,\mathfrak{v}^2\,.$$

Due to the formulas:

$$\tfrac{1}{2}\,\text{grad}\,\mathfrak{B}^2 = (\mathfrak{B}\ \text{grad}\ \mathfrak{B}) + [\mathfrak{B}\ \text{rot}\ \mathfrak{B}] \qquad \text{and} \qquad \frac{\partial\mathfrak{B}}{\partial t} + (\mathfrak{v}\ \text{grad})\ \mathfrak{B} = \frac{d\mathfrak{B}}{dt}\,,$$

from:

$$\text{rot}\ \mathfrak{v} = \frac{1}{\mu}\ \text{rot grad}\ S = 0\,,$$

one will ultimately get:

$$-\,\text{grad}\ U + \frac{h^2}{8\pi^2\mu^2}\,\text{grad}\,\frac{\Delta\alpha}{\alpha} = \mu\frac{d\mathfrak{v}}{dt}\,, \tag{20}$$

which can be written in the form $\mathfrak{K}_a + \mathfrak{K}_i = \mathfrak{K}$. Whereas $\partial\mathfrak{v} / \partial t$ means the change in velocity at a fixed spatial point, $d\mathfrak{v} / dt$ is the substantial acceleration of a point that moves with the fluid. On the left-hand side, $-$ (grad U) / μ corresponds to an *external force* $\mathfrak{K}_a$ per unit fluid mass, and $\frac{h^2}{8\pi^2\mu^2}\cdot\frac{\Delta\alpha}{\alpha}$ represents a force function of the *internal forces* $\mathfrak{K}_i$ per unit mass, whose gradient contributes to the acceleration of the points of the continuum. When the Ansatz (18) is introduced into equation (14), it will show that the total amount $\int \mu \alpha^2 dv$ of fluid will remain continually constant.

One can already see the necessity of introducing that *internal force* in the following example: A hydrogen atom is found in an energy state E_n . As will be established later, $|\psi_n|^2\, dV$ will then mean the probability that the electron will be found in precisely the volume *dV* with a total energy of E_n . If *dV* lies at a distance *r* from the nucleus then for a sufficiently-large *r*, the potential energy $-\varepsilon^2 / r$ will be more weakly negative than the given total energy E_n such that only a *negative* residual magnitude will remain for the kinetic energy $\frac{1}{2}\mu\mathfrak{v}^2$, which will belong to an *imaginary* velocity $\mathfrak{v}$, i.e., the electron will be found in *dV* with a probability of $\psi_n^2\, dV$, but with an imaginary velocity, which is naturally absurd. The solution to the paradox lies in just the fact that one cannot interpret the negative residual magnitude as kinetic energy alone. The same reason for saying that $|\psi_n|^2\, dV$ can be interpreted as the occupation density of the electron in *dV* will turn (17) into:

$$\frac{h}{4i\pi\mu}\left(\frac{\operatorname{grad}\tilde{\psi}}{\psi} - \frac{\operatorname{grad}\psi}{\tilde{\psi}}\right) = \mathfrak{v}$$

as the expression for the *real* velocity of the electron in *dV*. However, that will be mechanically possible only when the potential energy $-\ \varepsilon^2 / r$ has a suitable negative internal potential energy U_i added to it, which is taken to be large enough that the energy balance:

$$E_n = -\ \frac{\varepsilon^2}{r} + U_i + \tfrac{1}{2}\mu\mathfrak{v}^2$$

will be true.

In special case of a *time-constant* potential function $U_0\,(q)$, ψ will be equal to, e.g., $a_n \cdot \psi_n e^{-2i\pi E_n t/h}$ with *time-constant* coefficients.

$$\rho = \alpha^2 = \psi\tilde{\psi} = a_n^2 \psi_n^2 (q)$$

is then *time-constant* in any event, i.e., one is dealing with a *stationary* state of flow that is even static ($\mathfrak{v} = \frac{1}{\mu}$ grad $S = 0$), moreover, since (18′) will be:

$$S = \frac{h}{4i\pi}(\ln\psi - \ln\tilde{\psi}) = -\,E_n\, t$$

here and have vanishing gradients.

Apart from its greater intuitive appeal, the hydrodynamical interpretation of the function $\psi(q,t)$ has yet another special significance due to the fact that it suggests that one can couple the wave-mechanical state of an atom with the surrounding electrodynamical field by regarding the hydrodynamical charge and current densities $\varepsilon \rho$ and εj as the sources and points of application of electrodynamical fields according to **Maxwell**'s theory. However, one must notice immediately that one will contradict experiments in that way, because in the special case of $| a_n |^2 = 1$, $a_k = 0$ for $k \neq n$ (atom in the state n) with $\psi = \psi_n(q) \cdot e^{-2i\pi E_n t/h}$, (17) will give $j = 0$, and in that way according to **Maxwell**'s theory, there will be no radiation, whereas in reality, an atom in the state n will emit a ray spontaneously.

However, there are other ways of seeing that the above values of ρ and j that are derived for the state ψ [for the state of $N_1 = | a_1(t) |^2$ atoms in the state 1, etc., with the statistical picture] have nothing to do with the emission and absorption of radiation by that collection of atoms. That is because if one imagines, say, all N atoms being pushed together without mutually perturbing each other, then no forces will appear between the parts of different atoms, just as in the separated state. The values of j and ρ that were calculated by the method above will remain unchanged under that compaction, while it is known that the **Maxwell** radiation that is emitted from separated systems of atomic electrons is completely different from that of compacted systems of electrons. We shall take that lack of utility of the continuous interpretation of the ψ-function for calculating the interaction with radiation to be a good reason for later replacing the continuous interpretation with another one, namely, the *statistical interpretation*.

It is instructive to also present the argument above with the influence of a *magnetic field*. According to no. **51**, eq. (6), the **Schrödinger** equation will then read:

$$\Delta \psi - \frac{8\pi^2 \mu}{h^2} \varepsilon \varphi \cdot \psi - \frac{4i\pi\mu}{h} \left(\frac{\partial \psi}{\partial t} - \frac{\varepsilon}{c\mu} \mathfrak{A} \operatorname{grad} \psi \right) = 0 , \tag{21}$$

with φ as the scalar potential and $\mathfrak{A}$ as the vector potential that act upon the charge ε. A calculation that is analogous to one above, along with the use of div $\mathfrak{A} = 0$ and the Ansatz $\psi = \alpha e^{2i\pi S/h}$, will lead to:

$$\left. \begin{aligned} &\rho = \alpha^2 = \text{density}, \quad j = \frac{\alpha^2}{\mu} \left(\operatorname{grad} S + \frac{\varepsilon}{c} \mathfrak{A} \right) = \text{current density}, \\ &\mathfrak{v} = \frac{j}{\rho} = \frac{1}{\mu} \left(\operatorname{grad} S + \frac{\varepsilon}{c} \mathfrak{A} \right) = \text{current density}, \end{aligned} \right\} \tag{22}$$

and to the equation of motion:

$$\mu \frac{d\mathfrak{v}}{dt} = \varepsilon \left(-\operatorname{grad} \varphi - \frac{1}{c} \frac{\partial \mathfrak{A}}{\partial t} \right) + \frac{\varepsilon}{c} [\mathfrak{v}, \operatorname{rot} \mathfrak{A}] + \frac{h^2}{8\pi^2 \mu} \operatorname{grad} \frac{\Delta \alpha}{\alpha} , \tag{22′}$$

when one consistently drops relativistic terms with $\mathfrak{A}^2 / c^2$. Due to the facts that:

$$\mathfrak{E} = -\operatorname{grad}\varphi - \frac{1}{c}\dot{\mathfrak{A}}, \qquad \mathfrak{H} = \operatorname{rot}\mathfrak{A}, \qquad \mathfrak{K}_i = \frac{h^2}{8\pi^2\mu}\operatorname{grad}\frac{\Delta\alpha}{\alpha}, \tag{23}$$

an "internal force" $\mathfrak{K}_i$ will be added to the usual force $\mathfrak{K}_a$ of the external field.

38. Wave packets (1). – If one has a solution $\psi(q, t)$ for which $\psi\tilde{\psi} = \rho$ possesses considerable values only in a small neighborhood of a coordinate point P_0 at a time point t_0, but relatively-small values at some distance from P_0, then one speaks of a "wave packet" that is concentrated into a small space around P_0 . If t increases then one will generally find, first of all, a change in the location of the maximum of ρ and secondly, a broadening that can lead to a complete dissipation of the maximum over the course of time. Now, the change in location of the wave packet has a strong similarity to the motion of a mass-point in classical mechanics. Namely, if one multiplies equation (20):

$$\frac{d}{dt}(\mu\mathfrak{v}) = -\operatorname{grad} U + \frac{h^2}{8\pi^2\mu}\operatorname{grad}\left(\frac{\Delta\alpha}{\alpha}\right)$$

by $\rho = \alpha^2$ and integrates over the entire space dV then the integral over the *last* term on the right will drop out with the help of partial integrations (2), and what will remain is:

$$\int \frac{d}{dt}(\mu\mathfrak{v})\cdot dV = \int \rho\cdot(-\operatorname{grad} U)\, dV .$$

One can write that equation in the form:

$$\overline{\frac{d}{dt}(\mu\mathfrak{v})} = \overline{\mathfrak{K}_a} ,$$

---

(1) **E. Schrödinger**, Ann. Phys. (Leipzig) **79** (1926), pp. 489.

(2) The x-component is:

$$\begin{aligned}\alpha^2 \operatorname{grad}_x \frac{\Delta\alpha}{\alpha} &= \alpha \operatorname{grad}_x \Delta\alpha - \Delta\alpha\cdot\operatorname{grad}_x \alpha = \operatorname{grad}_x(\alpha\,\Delta\alpha) - 2\,\Delta\alpha\cdot\operatorname{grad}_x\alpha \\ &= \operatorname{grad}_x(\alpha\,\Delta\alpha) - 2\,\{\operatorname{div}(\operatorname{grad}_x\alpha\cdot\operatorname{grad}\alpha) - \operatorname{grad}\alpha\cdot\operatorname{grad}\operatorname{grad}_x\alpha\} \\ &= \operatorname{grad}_x(\alpha\,\Delta\alpha) - 2\,\{\operatorname{div}(\operatorname{grad}_x\alpha\cdot\operatorname{grad}\alpha) + \operatorname{grad}_x(\alpha\cdot\operatorname{grad}\alpha)^2 - 2\,[\operatorname{grad}\alpha, \operatorname{rot}\operatorname{grad}\alpha]_x .\end{aligned}$$

The last term vanishes here because rot grad = 0, and what will remain is:

$$\alpha^2 \operatorname{grad}_x \frac{\Delta\alpha}{\alpha} = \frac{\partial}{\partial x}[\alpha\,\Delta\alpha + (\operatorname{grad}\alpha)^2] - 2\operatorname{div}(\operatorname{grad}_x\alpha\cdot\operatorname{grad}\alpha) .$$

If the right-hand side is integrated over all of space with $dx\,dy\,dz$ then that will give zero when α vanishes strongly enough at infinity, along with its derivatives.

with the mean values of the acceleration and force:

$$\overline{\frac{d\mathfrak{v}}{dt}} = \int \frac{d\mathfrak{v}}{dt} \cdot \rho \, dV \,, \qquad \overline{\mathfrak{K}_a} = \int (-\operatorname{grad} U) \cdot \rho \, dV \,, \qquad \overline{\mathfrak{K}_i} = \int \frac{h^2}{8\pi^2 \mu} \alpha^2 \operatorname{grad} \frac{\Delta\alpha}{\alpha} dV = 0 \,,$$

i.e., the center of mass of the wave packet moves as if only the *external* force acts upon the entire wave packet, while the *internal* force cancels out completely. If the wave packet is densely concentrated around its center of mass then the result that was just found can also be expressed: *The acceleration of the center of mass of the wave packet is consistent with an external force* $(-\operatorname{grad} U)$ *that prevails at the location of the wave packet in the sense of* ***Newton****'s equations.* According to that law, which **Ehrenfest** (1) found and is still true in the absence of a magnetic field (23), it would be very tempting to relate that to the original **Schrödinger** viewpoint that material particles, electrons, etc., are nothing but wave packets. However, that would contradict the fact that wave packets generally dissipate in the course of time. In order to pursue that effect, we consider electrons (2) that are flowing freely in the x-direction with the Hamiltonian function $H(p, q) = \frac{1}{2\mu} p_x^2$, so it will have the wave equation:

$$\frac{\partial \psi}{\partial t} = a^2 \frac{\partial^2 \psi}{\partial x^2} \qquad \text{with} \qquad a^2 = \frac{i h}{4\pi \mu} .$$

That has the form of the *equation of heat conduction*. We shall consider its general solution:

$$\psi(x, t) = \frac{1}{2a\sqrt{\pi t}} \int_{-\infty}^{+\infty} d\xi \cdot e^{-\frac{(x-\xi)^2}{4a^2 t}} \cdot \psi(0, \xi) \tag{24}$$

for the following special form of the initial state:

$$\psi(x, 0) = C \cdot e^{-x^2/2\omega^2 + i\alpha x}$$

(ω and α are arbitrary real constants), so:

$$\rho(x, 0) = \psi(x, 0) \cdot \tilde{\psi}(x, 0) = C^2 \cdot e^{-x^2/\omega^2} .$$

The initial state represents a wave packet with a maximum that gets steeper as ω gets smaller. After some calculations, the general solution (24) then leads to the following density distribution for the wave packet at time t:

---

(1) **P. Ehrenfest**, Zeit. Phys. **45** (1927), pp. 455.

(2) Some more-complicated cases of electrons in electric and magnetic fields were treated by **C. G. Darwin**, Proc. Roy. Soc. **117** (1927), pp. 258. Furthermore, see **Kennard**, Zeit. Phys. **44** (1927), pp. 326.

$$\psi\tilde{\psi} = c(t)\cdot e^{-\left(x-\frac{h\alpha}{2\pi\mu}t\right)^2/\Omega^2},$$

in which:

$$\Omega^2 = \omega^2 + \frac{h^2\,t^2}{4\pi^2\mu^2\,\omega^2},$$

i.e., a displacement of the wave packet with the velocity $h\ \alpha\ /\ 2\pi\,\mu$ that flattens out as time increases. In particular, the initial width (i.e., $\Omega^2 = 4\omega^2$) will *double* after a length of time equal to:

$$t = \sqrt{3}\cdot\frac{2\pi\,\mu\,\omega^2}{h}.$$

For $\mu = 1.7 \cdot 10^{-24}$ g (*H*-atom) and $\omega = 10^{-8}$ cm (atomic diameter), one has $t \approx 10^{-13}$ sec. By contrast, for $\mu = 10^{-3}$ g and $\omega = 10^{-3}$ cm, one will have $t \approx 10^{18}$ sec = $\frac{1}{3}\cdot 10^{11}$ years. Due to that tremendous length of time, the idea that macroscopic masses might be wave packets does not initially contradict anything from an experimental standpoint. On the other hand, experiments also show the existence of "microscopic" particles (electrons, α-particles, etc.) that do not dissipate, no matter how old they get.

For that reason, we will regard the hydrodynamical interpretation of the ψ-field as a picture that is useful for only imagining the properties of the ψ-function, but we shall not speak of the hydrodynamical quantities ρ and j as the density and current density of real matter itself. Rather, in what follows, we shall pursue an entirely different sort of *statistical interpretation* of the ψ-field.

Before we treat it systematically in nos. **39**, *et seq.*, we would like to illustrate it here in the example of the wave packet. Therefore, we have a ψ-function whose magnitude $|\,\psi\,|^2$ at time t_0 has a value that is appreciable only in the close neighborhood of a coordinate point P_0. The value of $|\,\psi\,|^2\,dv$ at time t_0 will then be interpreted as the *probability* that the system exists in precisely the volume element dV at the point P, and indeed let that probability be appreciable only close to P_0 at just that time $t = 0$. However, the probability density $|\,\psi\,|^2$ for any later or earlier time t can now be calculated *inevitably* (i.e., causally) from that *arbitrarily* given initial probability density in coordinate space v with the help of **Schrödinger**'s spacetime differential equation for ψ. One can also express that as follows: One gives initial positions to a large number of systems of the same type such that the initial number of systems in the region dv will be proportional $|\,\psi\,|^2\,dv$. The statistical distribution $|\,\psi\,|^2$ of the ensemble of systems at a later or earlier time t will then be governed by **Schrödinger**'s equation for ψ. However, that says absolutely nothing about the individual fate of an isolated system in the ensemble. That is connected with the fact that the *energies* and *velocities* of the individual systems of the ensemble at time $t = 0$ are not at all determined for our spatial distribution of the system at the time t_0. In contrast to that, the sharper the peak of $|\,\psi\,|^2$ at the time t_0 and at the point P_0, i.e., the sharper the initial position of the system is constricted, the more that the various eigenfunctions ψ_1, ψ_2, … in the eigenfunction development:

$$\psi(q, t) = \sum_k c_k\, e^{2i\pi E_k/h}\, \psi_k(q)$$

will be assigned appreciable coefficients $c_k(t_0)$ in the construction of ψ, i.e., the more that the system of the ensemble already belongs to the various energy and impulse values with frequencies $c_1^2(t_0) : c_2^2(t_0) : c_3^2(t_0) : \ldots$ at time t_0 . Conversely, if only systems in the *one* energy state E_j were present at time t_0 then their spatial distribution would be given by the density function $|\,\psi_j(q)\,|^2$, which does not define a sharp peak, but a broad spatial region for the initial spatial positions of the individual systems. (For the reciprocity of the precision of the values of the coordinates and velocities, or time and energy, cf. no. **19**.)

Now, every more or less *sharp* initial distribution of the *positions* of the individual systems, i.e., every more or less *indistinct* initial distribution of the *velocities* of the individual system, belongs to a well-defined initial position and initial velocity of the *center of mass* of the system distribution. The law of **Ehrenfest** that was proved (cf., *supra*) says that the *center of mass* advances as an isolated system that is given its initial position and initial velocity would advance according to *classical* mechanics.

The rapid broadening of an initially sharp spatial system distribution is attributed to the associated large uncertainty in the initial velocity distribution. Conversely, for an initial distribution over a large spatial region, the associated initial velocities of the system points will be less scattered, and the broadening of the spatial region will correspondingly take place more slowly. It hardly needs to be mentioned that it would completely contradict the conception of things that was just presented if one regarded a wave packet with a certain initial spatial extent as an aggregate of extended matter (massive object) and then hoped to glimpse an image of a spatial blurring of the massive object in the broadening of the wave packet in time. In contrast, the statistical interpretation is based upon the picture of each individual system occupying a sharply-defined coordinate point at each moment and the idea that the statistical distribution of the point positions at time t can be calculated from a given statistical distribution of many such spatial points at time t_0 .

39. Field theory and corpuscular theory. – We might now go into the duality between the wave and corpuscular theories more generally.

As is known, the wave theory of *light*, as opposed to the older seemingly-surpassed emission theory, has gained a new opponent in the form of ***Einstein****'s theory of light quanta*, which is built upon the fact that many optical observations are not consistent with the picture of a continuous spreading of light energy and impulse. Here, one recalls the photoelectric effect, i.e., release of individual electrons from the individual locations on irradiated metal by incident light long *before* sufficient light energy in the form of continuous waves has been achieved at those locations. Moreover, the energy E of the emitted electrons does not depend upon the intensity, but on the *color* ν of the light ($E \leq h\nu$). That was further verified by the experiments of **Bothe** and **Geiger**, which showed a direct coupling of a process of atomic emission or light scattering with a simultaneous process of absorption of finite energy. That and many other phenomena support the **Einstein** hypothesis that light does not propagate continuously, but in the form of corpuscular light quanta whose energy and impulse are connected with the frequency ν and wavelength λ that might present themselves by the relations $E = h\ \nu$, $p = h / \lambda$. Of course, in that way, due to the fact of the interference and wave nature of the radiation that is supported by the latter, the light corpuscles can likewise be ascribed a higher degree of reality, as one was accustomed to do for material corpuscles, such as the indestructible electrons and protons of cathode ray and canal ray experiments (but has abandoned in the meantime). The attempt to regard light quanta as a narrowly-confined aggregate of waves that comes about by the superposition of suitable solutions to **Maxwell**'s equations could never get to the heart of the matter, first of all, due to the appearance of the quantity h, which is alien to **Maxwell**'s equations, and furthermore, due to the fact that when a narrowly-confined wave-aggregate impinges upon something with an inhomogeneous index of refraction, e.g., a semi-transparent mirror, a scattering or *division* of the wave-aggregate must occur, which contradicts the experiments that have led to the assumption of the indivisibility of light quanta.

Exactly the same thing is true of the theory of *matter*. Here, supported by the results of the kinetic theory of gases and the study of electrons, the corpuscular nature of matter seems to have triumphed over the older continuum theories. However, quantum mechanics, in the form that **Schrödinger** gave it, is compelled by the wealth of quantum-mechanical phenomena to juxtapose an undulatory theory of matter with point mechanics. However, complications will appear in that way that correspond completely to the ones in the optical case, since the assumption that point-like electrons and the systems that are constructed from them, such as atoms and molecules, are indestructible corpuscles is so convincing (one imagines, say, the trajectory of an α-particle in the **Wilson** cloud chamber) that it would seem difficult, with no further analysis, to go over to an undulatory theory of matter in which that indestructibility and distinguishability of the particle should no longer be true, at least to a certain degree.

Now it is, of course, possible to define wave-aggregate withing the context of wave mechanics that behave like mass-points in many respects, since that is just what the optical-mechanical analogy shows. For example, one can (no. **38**) construct a wave-packet that occupies a very narrow space and moves like a classical mass-point from a superposition of solutions to the equation of undulation with suitable phases and amplitudes (just as light quanta can be approximated by wave bundles). However, if one gives a greater or lesser degree of inhomogeneity to the "index of refraction" that the wave-aggregate encounters at some location, i.e., a non-constant or even discontinuous behavior of the potential at that location, then the wave-aggregate will be *scattered* or divided there and will no longer have the character of a mass-point. However, according to experiments, an electric particle will still be recognized to be unchanged as a mass-point after a collision. The undulatory conception of mechanics will then give *only one* aspect of the experiment in any event.

The same duality would exist here that exists in optics then if it were not for the fact that there is an overarching concept here as well as there. It consists of the *statistical* interpretation of the wave-theoretic results in the sense that the wave function ψ is referred to as a *measure of the probability* that corpuscular particles are found at the individual locations in space and time.

The statistical picture shall be discussed in more precise detail in the following sections on the basis of the fundamental equation of wave mechanics and its solutions.

40. Statistical interpretation of the wave function. – With the use of the coordinates q and t and the conjugate impulse p and $-E$, a quantum-mechanical system $H(q, t; p)$ will have the wave equation:

$$\{H(q, t; p) - E, \psi(q, t)\} = 0 \qquad \text{with} \qquad -E = +\frac{h}{2i\pi}\frac{\partial}{\partial t}, \qquad p_\kappa = \frac{h}{2i\pi}\frac{\partial}{\partial q_\kappa}.$$

One asks what the physical interpretation of a given solution $\psi(q, t)$ would be when one keeps to the picture in which the mechanical system $H(q, t; p)$ is *kinematically* related to a *point-mechanical* system, i.e., it can be found in *only one* configuration at one moment, but not, say, fill up a whole region dv with a certain density at time t. Dynamically, however, the succession of different configurations q shall not be governed by classical mechanics, but by quantum laws of a statistical type. According to an idea that was first expressed quite clearly by **Dirac** and **Born** (^{1}), and then successfully pursued by **Jordan** (^{2}), $\psi(q, t)$ should be interpreted as a *probability amplitude* and $|\psi|^2$ as the *probability* that an atom that is described by the Hamiltonian function $H(q, t; p)$ will be found in the region dv of coordinate space at time t. Assume that $\psi(q, t)$ is normalized to 1, i.e., $\int \psi\tilde{\psi}\, dv = 1$. The fact that this is possible follows from the constancy of $\int \psi\tilde{\psi}\, dv$ that is assured by (14).

Example 1. – Consider, in particular, the normalized eigen-solution:

$$\psi(q, t) = \psi_n(q) e^{-2i\pi E_n t/h}$$

for an eigenvalue E_n for a conservative system. Here, according to the definition that was just given:

$$\psi\tilde{\psi}\, dv = \psi_n\tilde{\psi}_n\, dv = |\psi_n(q)|^2\, dv$$

is the *probability* that the coordinates of the atom lie in dv for an energy value of E_n. Statistically, that means that in the presence of many mutually-independent atoms of the same type that are all in the state n, the relative number of systems that are found in dv will be equal to $\left|\psi_n^2(q)\right| dv$.

Example 2. – Consider the general solution:

$$\psi(q, t) = \sum a_n(t)\psi_n(q)\cdot e^{-2i\pi E_n t/h} \tag{25}$$

(^{1}) **P. A. M. Dirac**, Proc. Roy. Soc. **113** (1926), pp. 621; **M. Born**, Zeit. Phys. **38** (1926), pp. 803.

(^{2}) **P. Jordan**, Naturwiss. **15** (1927), pp. 105; Zeit. Phys. **40** (1927), pp. 661 and 809.

that is constructed from the normalized eigenfunctions $\psi_n (q)$ and time-dependent quantities $a_n (t)$ for a non-conservative system that might emerge from a conservative system with the eigen-solutions $\psi_n (q)$ by a perturbation.

We then assume that ψ is normalized to 1: $\int \psi \tilde{\psi} \, dv = 1$. We can then interpret $\psi \tilde{\psi} \, dv$ as the probability that the particle occupies dv. The *total probability* that particles exist somewhere in space is:

$$1 = \int \psi \tilde{\psi} \, dv = \sum |a_n|^2$$

[the latter is similar to (16′)]. It is composed additively from the $|a_n|^2$, which then refer to the partial probabilities that the particle is somewhat in space, but occupies the state n, in particular. Finally, one would like to interpret:

$$|a_n|^2 \, |\psi_n (q)|^2 \, dv$$

as the probability that the particle is in the state n and occupies the volume dv. However, in truth, the probability that particles enter into any state, and in so doing dv, is equal to, say, $\sum_n |a_n|^2 |\psi_n|^2 \, dv$, but from (25), it is equal to:

$$\psi \tilde{\psi} \, dv = \sum |a_n|^2 |\psi_n|^2 \, dv + \sum \sum a_n a_m \psi_n \tilde{\psi}_m \, e^{-2i\pi(E_n - E_m)t/h} dv,$$

which has an additional *interference term*. A *superposition* enters in place of the simple addition of the partial probabilities, since the *probability amplitudes* $a_n \psi_n$ with the phase factors $e^{-2i\pi E_n t/h}$ combine additively with the total amplitude ψ, and it is only the square of the latter that gives the *intensity* $\psi \tilde{\psi}$ of the probability itself. That interference of probabilities exhibits the peculiarity of quantum theory as undulatory mechanics in a characteristic way, and as **Jordan** (1) had shown, one can, conversely, start from the postulate of the interference of probabilities and successfully reconstruct quantum theory.

The probability-theoretic interpretation of the existence of *one* atom in the state n and volume element dv is equivalent to the following *statistical* interpretation: Let a very large number of mutually-independent atoms of the same type be present. The relative number of atoms that are found in the state n in an arbitrary configuration is $|a_n(t)|^2$, i.e., it behaves like:

$$N_1 : N_2 : \ldots = |a_1^2| : |a_2^2| : \ldots,$$

in which:

$$N_1 + N_2 + \ldots = N = \text{const.} \qquad \text{(conservation of particle number)},$$

because $\sum |a_n|^2$ is constant in time.

If one were to use a function that is not normalized to 1, but "normalized to N":

---

(1) **P. Jordan**, Zeit. Phys. **40** (1927), pps. 661 and 809.

$$\psi(q, t) = \sum c_n(t)\psi_n(q)\cdot e^{-2i\pi E_n t/h} \qquad \text{with} \qquad \sum c_n^2 = N ,$$

then $c_n^2(t)$ would be given directly by the number N_n of systems in the state n.

Example 3 (which will be referred to in no. **45**). – Consider a wave function $\psi(N_1, N_2, \ldots)$ that depends upon the "coordinates" N_1, N_2, … and is the solution of a certain transformed wave equation. We once more interpret ψ as the probability amplitude and $|\psi(N_1, N_2, \ldots)|^2$ as the probability that the system occupies the "coordinate point" N_1, N_2, … $|\psi(N_1', N_2', \ldots)|^2$ correspondingly gives the probability that the system will occupy the coordinate point N_1', N_2', ... The "coordinates" N_1, N_2, … that characterize the "position" of the system might have the following meaning: The total system consists of a number of mutually-independent subsystems of the same type, of which N_1 are in the state 1, N_2 are in the state 2, etc. For a large ensemble of total systems $|\psi(N_1, N_2, \ldots)|^2$ will then give the relative number of total systems that are found in precisely the state N_1, N_2, …, i.e., their subsystems are distributed over the states 1, 2, … with the numbers N_1, N_2, … In this example, we see that the statistical interpretation of ψ can will be quite different according to the meaning of the "coordinates."

Whereas the normalized eigen-solutions $\psi_n(q)$ of a conservative system are fixed once and for all, the series $\psi(q, t) = \sum a_n \psi_n$ can still have a very different form according to the form of the functions $a_n(t)$. Now, wave mechanics determines the $a_n(t)$ only to the extent that it *fixes the time evolution of the* $a_n(t)$ *for arbitrarily-given initial values* $a_n(t_0)$. That is, one assumes the distribution of a distribution of many independent atoms over the state n at time t when the initial distribution is given. By contrast, that does not determine the fate of an individual system that is now found in precisely the state n after a certain amount of time t has evolved. Wave mechanics then includes no causal determination of the states that follow each other from an individual isolated system. Rather, it only calculates the *probability* that the isolated system in a given initial state will reach new states.

If one perceives the quantum mechanics of electrons to be the final regulator of all things (for the time being, there is no reason for assuming that there are further hidden mechanisms that might be at work) then one would have to give up on the causal determinism of classical theory in favor of only a statistical determinism of physical events. For macroscopic observations, exact causality will then present itself as asymptotically statistically legitimate then. However, the more detailed the microscopic relationships are considered, the more strongly that uncertainties and dispersions will play a role, since they have their empirical counterparts in the chiefly quantum-theoretic fuzziness of observation.

41. Forced quantum transitions (1). – Let the effect of a perturbing potential $F(q, t)$ be imposed upon a mechanical system that is described by the Hamiltonian function $H^{\mathrm{o}}(q, p) = T + U^{\mathrm{o}}(q)$ when it is unperturbed. The wave equation of the system, with $H = H^{\mathrm{o}} + F$ is then:

$$\left\{H + \frac{h}{2i\pi}\frac{\partial}{\partial t}, \psi\right\} = \left\{H^{\mathrm{o}}\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + F(q,t) + \frac{h}{2i\pi}\frac{\partial}{\partial t}, \psi(q,t)\right\} = 0\,. \tag{26}$$

We take a solution to that equation to be a series in the eigenfunctions ψ_n of the unperturbed equation:

$$\left.\begin{aligned} &\psi(q,t) = \sum_m b_m(t)\,\psi_m(q) = \sum_m a_m(t)\,e^{-2i\pi E_m t/h}\;\psi_m(q)\,, \\ &\text{with} \\ &b_m(t) = a_m(t)\,e^{-2i\pi E_m t/h}. \end{aligned}\right\} \tag{26′}$$

In that way, $b_m(t)$ includes the phase factor $e^{-2i\pi E_m t/h}$, which is *rapidly-varying in time*, while $a_m(t)$ varies only slowly with t. (The *phase* is $\Theta_m = E_m t : h$.)

According to no. **40**:

$$|\,b_m(t)\,|^2 = |\,a_m(t)\,|^2 = N_m(t)$$

means the *number* of atoms in the state n. Instead of asking how that number would change in time under the influence of the perturbing field directly, we first examine the simpler question of how the $a_m(t)$ or $b_m(t)$ would change in time ($db_m / dt = b'_m$).

If one substitutes the series (24′) then one will get:

$$\begin{aligned} 0 &= \sum_m b_m \cdot \{H^{\mathrm{o}}\,\psi_m\} + F \cdot \sum_m b_m\,\psi_m + \sum_m \psi_m\,\frac{h}{2i\pi}\,b'_m \\ &= \sum_m b_m\,E_m\,\psi_m + F\sum_m b_m\,\psi_m + \frac{h}{2i\pi}\sum_m b'_m\,\psi_m\,. \end{aligned} \tag{26″}$$

One develops the $F\,\psi_m$ in that into an eigenfunction series:

$$F(q, t)\,\psi_m(q) = \sum_k F_{km}(t)\,\psi_m(q)\,, \tag{27}$$

whose coefficients $F_{km}(t)$ are determined to be:

$$F_{km}(t) = \int \tilde{\psi}_k(q)\,F(q,t)\,\psi_m(q)\,dv = \tilde{F}_{mk}(t) \tag{27′}$$

---

(1) **M. Born**, "Das Adiabatenprinzip in der Quantemechanik," Zeit. Phys. **40** (1926), pp. 167.

due to the orthogonality and normalization of the $\psi_k(q)$. One calls F_{km} a *matrix element* (1) of $F(q)$. The splitting of $H(q, p, t)$ into $H^0(q, p) + F(q, t)$ in that shall be done in such a way that the matrix elements F_{nn} vanish when the part of the perturbing energy that does not give rise to vanishing F_{nn} is attributed to the principal energy H°. One will then have:

$$H_{km} = F_{km} \qquad \text{for} \qquad m \neq k, \qquad\qquad H_{kk} = E_k\,, \qquad F_{kk} = 0\,. \tag{27''}$$

Substituting the series (27) in (26″) will give:

$$0 = \sum_k \psi_k \cdot \left[b_k\, E_k + \sum_m F_{km}\, b_m + \frac{h}{2i\pi}\, b_k' \right].$$

Multiplying that equation by a certain ψ_k and integrating over dv will ultimately give:

$$0 = b_k\, E_k + \sum_m F_{km}\, b_m + \frac{h}{2i\pi}\, b_k' \qquad\qquad \text{for} \qquad k = 1, 2, \ldots,$$

for which we write:

$$-\frac{h}{2i\pi} b_k' = \sum_m H_{km}(t)\, b_m(t) \quad \text{with} \quad H_{km} = \begin{cases} F_{km} & \text{for} \quad m \neq k \\ E_k + F_{kk} & \phantom{\text{for}} \quad m = k\,. \end{cases} \tag{28}$$

That is the law for that temporal change in the b_k. If one introduces equation (26′) into that then one will get the following equation for the slowly-varying a_k:

$$-\frac{h}{2i\pi} a_k' = \sum_m F_{km}(t) \cdot a_m(t)\, e^{2i\pi(E_k - E_m)t/h}. \tag{28'}$$

We contrast the last two equations with the starting equation (26) in the form:

$$-\frac{h}{2i\pi} \dot{\psi} = \{H,\, \psi\}. \tag{29}$$

---

(1) In general, the matrix element of a function $A(q, p)$ relative to the eigenfunctions ψ_m of a Hamiltonian function will refer to the quantities $A_{km} = \int \tilde{\psi}_k(q)\, A\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) \psi_m(q)\, dv$, where each ψ is possibly provided with a time factor, such that:

$$\frac{dA_{km}}{dt} = \left(\frac{2i\pi}{h}\right)(\nu_k - \nu_m) A_{km}\,.$$

(For matrix algebra, see no. **65**, as well as no. **34**.)

Now, in order to pursue the time evolution of the a_k or b_k, equations (28′), (28) must be integrated. *In the first approximation*, for a small time interval from $t = 0$ to t, the slowly-varying quantities $a_m(t)$ on the right-hand side of (28′) can be treated as constants, and one will find that integration *in the first approximation* will give:

$$a_k(t) = a_k(0) + \sum_m f_{km}^{(1)}(t) \cdot a_m(0)\,, \tag{30}$$

with

$$f_{km}^{(1)}(t) = -\frac{2i\pi}{h} \int_0^t F_{km}(t) \cdot e^{2i\pi(E_k - E_m)t/h}\,, \tag{30′}$$

and in particular, $f_{kk}^{(1)} = 0$, due to (27″). The first approximation no longer suffices for large times, but the rigorous solution can also be represented by a series:

$$a_k(t) = a_k(0) + \sum_m f_{km}(t) \cdot a_m(0)\,, \tag{31}$$

in which the f_{km} now deviate from the quantities f_{km}^{1}, but can be given by a series development $f_{km} = f_{km}^{1} + \dots$ that **Born** gave.

One can best recognize the *physical meaning* of $f_{km}(t)$ in the special case where only one a [say, $a_l(0)$] is non-zero at time $t = 0$, i.e., all atoms are in the state l at time t. From (31), one will then have:

$$a_k(t) = f_{kl}(t) \cdot a_l(0)\,, \quad N_k(t) = |\,a_k(t)\,|^2 = |\,f_{kl}(t)\,|^2 \cdot N_l(0) = w_{kl}(t) \cdot N_l(0)$$

at time t. Here, one will refer to $|\,f_{kl}(t)\,|^2 = w_{kl}(t)$ as the *transition probability* that an atom will jump from the state l to the state k during the time t, and $|\,f_{kl}(t)\,|^2 N_l$ is the probability that N_l atoms will go from l to k. $f_{kl}(t)$ itself is then appropriately referred to as the *probability amplitude* for a transition $l \rightarrow k$ to occur during t.

In general, for arbitrary initial values $a_k(0)$, the $a_k(t)$ are coupled to the $a_k(0)$ according to (31) by the linear addition law:

$$a_k(t) - a_k(0) = \sum_m f_{km}(t) \cdot a_m(0)\,,$$

while the absolute value of the square is given by the law:

$$|\,a_k(t) - a_k(0)\,|^2 = \sum_m |\,f_{km}(t)\,|^2 \cdot |\,a_m(0)\,|^2 + \sum_m \sum_l f_{km}\, \tilde{f}_{kl}\, a_m(0)\, \tilde{a}_l(0)\,, \tag{32}$$

which one refers to as the *law of superposition*, and can write in the form [when $a_k(0) = 0$, $N_k(0) = 0$]:

$$N_k(t) = \sum_m w_{km}(t) \cdot N_m(0) + \sum_m \sum_l f_{km}\, \tilde{f}_{kl}\, a_m(0)\, \tilde{a}_l(0)\,. \tag{32$'$}$$

Along with the *principal terms* $w_{km}\, N_m\,(0)$, which are proportional to the numbers $N_m\,(0)$ in the states m, it also includes *interference terms*, each of which relates to the presence of *two* states m and l. The fact that the probabilities appear, not in a law of addition, but a law of superposition, is an *essential* fact in wave mechanics and can be referred to as the *interference of probabilities* (see also no. **40**).

For *potential functions* $F\,(q,\,t)$ *that vary slowly in time*, **Born** could derive the asymptotic vanishing of the transition probabilities from one state to another. The numbers N_k would then remain constant in time, as ***Ehrenfest****'s adiabatic theorem* demands.

42. Quantum transitions in a radiation field. – The time-dependent perturbation now consists of an electric field $\mathfrak{E}\,(t)$ of constant direction, in particular. One must then replace the perturbation energy $F\,(q,\,t)$ with:

$$F\,(q,\,t) = -\,\mathfrak{M}\,(q)\,\mathfrak{E}\,(t)\cos\,(\mathfrak{E},\,\mathfrak{M})\,, \tag{33}$$

in which $(\mathfrak{E},\,\mathfrak{M})$ is the time-constant angle between the direction of the field and the electrical moment vector $\mathfrak{M}\,(q)$ of the system of electrons in the configuration q. From the previous section, the transition probability w_{kn} during time t will then be, in the first approximation:

$$w_{kn}\,(t) = \left| f_{kn}^{(1)} \right|^2 = \left(\frac{2\pi}{h} \right)^2 |\,\mathfrak{M}\cos(\mathfrak{E},\mathfrak{M})\,|^2 \cdot \left| \int_0^t \mathfrak{E}(t)\, e^{-2i\pi\nu_{km} t}\, dt \right|^2, \tag{34}$$

when one uses (30′) and the definition (27′) of the *matrix element* of $\mathfrak{M}\,(q)\cos\,(\mathfrak{M},\,\mathfrak{E})$. We can see the meaning of the last factor in a *random radiation field* as follows: Let $\mathfrak{E}\,(t)$ be represented between times 0 and t_1 by the Fourier integral:

$$\mathfrak{E}\,(t) = \int_{-\infty}^{+\infty} \mathfrak{E}_\nu\, e^{2i\pi\nu t} d\nu\,, \tag{35}$$

whose coefficients $\mathfrak{E}_\nu$ are inversely calculated to be:

$$\mathfrak{E}_\nu = \int_0^{t_1} \mathfrak{E}(t)\, e^{-2i\pi\nu_{km} t} dt\,, \qquad \text{such that} \qquad \mathfrak{E}_{-\nu} = \tilde{\mathfrak{E}}_\nu\,, \tag{35$'$}$$

from which the identity will follow:

$$\int_{-\infty}^{+\infty} \mathfrak{E}_\nu \tilde{\mathfrak{E}}_\nu \, d\nu = \int_0^{t_1} \mathfrak{E}(t)\,dt \int_{-\infty}^{+\infty} \mathfrak{E}_\nu \, e^{-2i\pi\nu t} d\nu = \int_0^{t_1} \mathfrak{E}(t)\tilde{\mathfrak{E}}(t)\,dt ,$$

i.e.:

$$\int_{-\infty}^{+\infty} |\mathfrak{E}_\nu|^2 \, d\nu = \int_0^{t_1} |\mathfrak{E}(t)|^2 \, dt . \tag{35″}$$

Now, the *mean density* of the radiation energy is:

$$\rho = \frac{1}{4\pi} \overline{\mathfrak{E}(t)^2} = \frac{1}{4\pi t_1} \int_0^{t_1} |\mathfrak{E}(t)|^2 \, dt , \tag{36}$$

and due to (35″), one will have:

$$\rho = \frac{1}{4\pi t_1} \int_{-\infty}^{+\infty} |\mathfrak{E}_\nu|^2 \, d\nu = \int_0^{\infty} \rho_\nu \, d\nu , \quad \text{with} \quad \rho_\nu = \frac{1}{2\pi t_1} |\mathfrak{E}_\nu|^2 . \tag{36′}$$

Finally, with the use of (35′), the transition probability (34) to be calculated will then be:

$$w_{kn}(t) = \left(\frac{2\pi}{h}\right)^2 | \mathfrak{M} \cos(\mathfrak{E}, \mathfrak{M})_{kn} |^2 \, |\mathfrak{E}_\nu|^2 .$$

If one sets $|\mathfrak{E}_\nu|^2 = 2\pi t \cdot \rho(\nu_{kn})$ here, corresponding to (36′), and gives $\cos^2(\mathfrak{E}, \mathfrak{M})$ the value 1/3, in the event that the direction of $\mathfrak{E}$ fluctuates irregularly in all directions, then the desired transition probability during the time t will ultimately become:

$$w_{kn}(t) = \frac{8\pi^3}{3h^2} | \mathfrak{M}_{kn} |^2 \cdot \rho(\nu_{kn}) \cdot t \tag{37}$$

in which $\rho(\nu_{kn})$ gives a monochromatic section of the spectrum $\rho = \int_0^\infty \rho(\nu)\,d\nu$ of the random radiation, and $\mathfrak{M}_{kn}$ is defined as in (27′). One will also get the same value (37) for the inverse transition probability from k to n that is forced by the radiation field $\rho(\nu_{kn})$.

Along with the transition probabilities w_{kn} and w_{nk} that are *forced* by the radiation field $\rho(\nu_{kn})$, there are also *spontaneous* transition probabilities from n to k, where $E_n > E_k$, but not from k to n. One can include them in the calculations only when one considers the reaction of the electron system to the field. That will be done in Chapter V, where atoms and fields will be considered together as a quantum-mechanical system. The flaw in the method above, which does not lead to spontaneous radiation, consists of the fact that the electromagnetic field "consists of only light quanta," or better yet, it is coupled with the atom in a quantum way that contradicts the classical

Ansatz $F(q, t) = -(\mathfrak{M}, \mathfrak{E})$. The interaction between light and matter cannot be consistently achieved by translating **Hamilton**'s equation for a point-like atom into a *continuous* electromagnetic field, but rather one must start from a point-mechanical atom in a corpuscular light field and simultaneously make the transition to the wave theory of light and matter.

43. Spontaneous transitions. – As was shown in no. **40**, the function:

$$\psi(q, t) = \sum_k a_k(t) \cdot \psi_k(q) e^{-2i\pi E_k t/h} \tag{38}$$

can be used to characterize a state in which $|a_k|^2 = N_k$ atoms of an aggregate are found on the energy level E_k ($k = 1, 2, \ldots$). On the other hand:

$$\psi\tilde{\psi} = \sum_k \sum_l a_k \tilde{a}_l \psi_k \tilde{\psi}_l e^{2i\pi(E_l - E_k)t/h} = \rho(q, t) \tag{38′}$$

was seen to be a sort of *density* that statistically gave the *probability* that an atom would occupy a unit volume at the coordinate point q at the time t.

However, that density ρ and the associated current density j had nothing do with the absorption and emission of radiation by the ensemble of atoms directly, as was explained in no. **37**. Nonetheless, it might be mentioned here without proof in connection with formula (38′) how one has calculated the intensity of the spontaneous emission from an ensemble of atoms that are first in the state l and can then go to various lower states k.

One takes the "transition density per atom" to be the quantity:

$$\rho_{kl}(q, t) = \psi_k \tilde{\psi}_l e^{2i\pi(E_l - E_k)t/h}, \tag{39}$$

i.e., a term in the development (38′), but with $a_k \tilde{a}_l = 1$. In sec. **50**, the fact that the charge density $\varepsilon \cdot \rho_{lk}$ can function as the source of the spontaneous electrodynamical radiation per atom under the transition $l \to k$ will then be given a wave-mechanical basis. We remark that the temporal period of the "transition density":

$$\nu_{lk} = \frac{E_l - E_k}{h}$$

fulfills the ***Bohr** frequency condition* precisely. In that way, ν_{lk} seems to be the *beat frequency* between the frequency $\nu_l = E_l / h$ of the initial state and that $\nu_l = E_l / h$ of the final state.

The *intensity* of the radiation is calculated from the electric moment that belongs to the transition density in the following way:

If $X(q)$ means the x-component of the electron moment for the atom in question in the coordinate position $(q_1, q_2, \ldots)$ of its electron system, and if the various configurations are occupied with the relative frequency density $\rho_{lk}(q, t)$ then the mean x-moment will be:

$$\int X(q)\,\rho_{lk}(q,t)\,dv = \int X(q)\,\tilde{\psi}_l(q)\,\psi_k(q)\,dv\,e^{2i\pi\nu_{lk}t} = X_{lk}\,e^{2i\pi\nu_{lk}t}, \tag{40}$$

with the factor X_{lk} as the *amplitude* and ν_{lk} as the *frequency*. Now, from classical electrodynamics, N_l atoms whose moments oscillate with the amplitude X_{lk} and frequencies ν_{lk} would spontaneously radiate the energy:

$$\left.\begin{aligned} dE &= dt\cdot N_l\cdot\frac{64\pi^4}{3c^3}\nu_{lk}^4\cdot(X_{lk}^2+Y_{lk}^2+Z_{lk}^2)\\ &= dt\cdot N_l\cdot\frac{64\pi^4}{3c^3}\nu_{lk}^4\cdot\mathfrak{M}_{lk}^2.\end{aligned}\right\} \tag{41}$$

The same expression will also be derived consistently for the quantum-theoretical spontaneous radiation in no. **49**, eq. (36′).

If the basic expression $\rho_{lk}(q, t)$ is time-constant in a special case then (40) will also be time-constant. That would be possible only when X_{lk}, Y_{lk}, Z_{lk} vanish. The radiation (41) would then be equal to zero, i.e., for a time-constant ρ_{lk}, the radiative transition $l \to k$ would not take place (it would be "forbidden"). The same thing would also be true for the transitions $l \to k$ and $k \to l$ under the influence of incident light since they are also coupled with the existence of the moment amplitudes $\mathfrak{M}_{lk}^2 = X_{lk}^2 + Y_{lk}^2 + Z_{lk}^2$ (Chap. V).

According to (41), the energies per unit time that are radiated by N_l atoms in the state l under the transitions to the states k, m, n (i.e., the intensities) behave like:

$$J_{lk} : J_{lm} : J_{ln} = \nu_{lk}^4\,\mathfrak{M}_{lk}^2 : \nu_{lm}^4\,\mathfrak{M}_{lm}^2 : \nu_{ln}^4\,\mathfrak{M}_{ln}^2 : \cdots \tag{42}$$

If the frequencies ν_{lk}, ν_{lm}, ν_{ln} differ from each other only slightly then one will have the approximate *intensity ratios:*

$$J_{lk} : J_{lm} : J_{ln} = \mathfrak{M}_{lk}^2 : \mathfrak{M}_{lm}^2 : \mathfrak{M}_{ln}^2 \tag{42′}$$

for the spectral lines ν_{lk}, ν_{lm}, ν_{ln}, … under *spontaneous* transitions from l.

CHAPTER V

INTERACTION OF RADIATION AND MATTER

44. Operator calculus. – In Sec. IV, it was shown how one can go from the classical **Hamilton** energy equation of a system with N degrees of freedom $H(q_1, \ldots, q_n, t, p_1, \ldots, p_N) - E = 0$ to the associated wave-mechanical differential equation. One replaces the impulse p_k and $- E$ in the (suitably-symmetrized) Hamiltonian function with the operators:

$$p_k \to \frac{h}{2i\pi}\frac{\partial}{\partial q_k} \qquad \text{and} \qquad -E \to \frac{h}{2i\pi}\frac{\partial}{\partial t}, \tag{1}$$

and applies the operator thus-obtained:

$$H\left(q, t, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + \frac{h}{2i\pi}\frac{\partial}{\partial t} \quad \text{to} \quad \psi(q, t)\,.$$

The solution ψ will then correspond to certain physical properties of the quantum-mechanical systems.

In order to obtain solutions, it is often preferable to go from the originally-chosen coordinates and impulses q, p to new canonically-conjugate ones q', p' in which the solutions are easier to find. The Hamiltonian function $H'(q', t, p')$, when transformed into the new coordinates, will then lead to the new operator:

$$H'\left(q', t, \frac{h}{2i\pi}\frac{\partial}{\partial q'}\right) + \frac{h}{2i\pi}\frac{\partial}{\partial t} \quad \text{to} \quad \psi'(q', t)\,.$$

Naturally, the solutions ψ' are completely different functions from the original ones $\psi(q, t)$. However, our main interest is in the search for the quantities that can be derived from the ψ and ψ' that will possess precisely the same values for ψ that they do for ψ', so they will be *invariant* under whatever coordinate system q or q' we happen to been chosen as a basis. That is because it is only such *invariants* that are capable of representing physically-measurable quantities that are independent of the coordinate system that we might choose. One such invariant is, e.g., the *energy*, which has the eigenvalue E_n in wave mechanics. In fact, it can be proved (Chap. VII) that the eigenvalues of the equations:

$$\left\{H\left(q, t, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - E, \psi\right\} = 0 \qquad \text{and} \qquad \left\{H'\left(q', t, \frac{h}{2i\pi}\frac{\partial}{\partial q'}\right) - E, \psi'\right\} = 0$$

will be the same, while the eigenfunctions $\psi_n(q)$ and $\psi'_n(q')$ that belong to the eigenvalue E_n are completely different. The same thing will generally be true of the *matrix components* (pp. 92, footnote 2) of any function $F(q, p) = F'(q', p')$: It can be shown that the F_{kl} are equal to the F'_{kl}. Thus, they are capable of being interpreted physically since they are invariants.

Under the replacement $p_k \to \frac{h}{2i\pi}\frac{\partial}{\partial q_k}$, any *function* of the coordinates q and the canonically-conjugate impulses p will become an *operator* that can be applied to a function ψ of the coordinates.

The transition to new coordinates and impulses (operators, resp.) initially affects only the *operator* $H - E$ in the **Schrödinger** equation, and only secondarily affects the ψ-functions as solutions of it. It is therefore important to clarify the simplest rules for calculating with operators.

Due to the product rule of differentiation:

$$\frac{d}{dx}ab = a\frac{db}{dx} + b\frac{da}{dx},$$

one has:

$$\frac{h}{2i\pi}\frac{\partial}{\partial q_1}q_1\psi - q_1\frac{h}{2i\pi}\frac{\partial}{\partial q_1}\psi = \frac{h}{2i\pi}\cdot\psi, \tag{2}$$

while:

$$\frac{h}{2i\pi}\frac{\partial}{\partial q_1}q_2\psi - q_2\frac{h}{2i\pi}\frac{\partial}{\partial q_1}\psi = 0\cdot\psi. \tag{2'}$$

If one drops the function symbol ψ from those equations then one will get the operator equations:

$$\frac{h}{2i\pi}\frac{\partial}{\partial q_1}q_1 - q_1\frac{h}{2i\pi}\frac{\partial}{\partial q_1} = \frac{h}{2i\pi}, \qquad \text{while} \qquad \frac{h}{2i\pi}\frac{\partial}{\partial q_1}q_2 - q_2\frac{h}{2i\pi}\frac{\partial}{\partial q_1} = 0,$$

or ultimately, when one understands p_K to mean the operator $h / 2i\pi\,\partial / \partial q_K$:

$$p_1\,q_1 - q_1\,p_1 = \frac{h}{2i\pi}, \qquad \text{while} \qquad p_1\,q_2 - q_2\,p_1 = 0. \tag{3}$$

The last two expressions mean nothing but an abbreviated notation for equations (2), (2′). (3) states that the operator $p_1 = \frac{h}{2i\pi}\frac{\partial}{\partial q_1}$ will indeed *commute* with q_2 in the order of the product ($p_1\,q_2 = q_2\,p_1$), but p_1 does not commute with q_1, as is shown by (2), (2′), which are (3) when written out in detail. Thus, the product rule of commutation is not valid for a conjugate pair of coordinates and impulses (operators).

Upon reverting to the detailed notation (2), one will easily prove the following general rules for the commutation of products:

$$\left.\begin{array}{l} p_K\,q_L - q_L\,p_K = \begin{cases} \dfrac{h}{2i\pi} & \text{for} \quad K = L, \\ 0 & \qquad\; K \neq L, \end{cases} \\ p_K\,p_L - p_L\,p_K = 0, \quad q_K\,q_L - q_L\,q_K = 0, \\ \text{for} \quad p_K = \dfrac{h}{2i\pi}\dfrac{\partial}{\partial q_K}. \end{array}\right\} \tag{4}$$

Along with that, when x_1, x_2, x_3 are three functions that depend upon the $2N$ quantities $q_1, \ldots, q_N$, $p_1, \ldots, p_N$ arbitrarily, the laws of association and distribution for addition will be true:

$$(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3)\,, \qquad (x_1\,x_2)\,x_3 = x_1\,(x_2\,x_3)\,,$$

$$x_1\,(x_2 + x_3) = x_1\,x_2 + x_1\,x_3\,, \quad x_1 = 0 \quad \text{or} \quad x_2 = 0 \quad \text{when} \quad x_1\,x_2 = 0\,,$$

and more generally, the commutative law of addition:

$$x_1 + x_2 = x_2 + x_1\,.$$

One can then calculate with the quantities q_K and the operators $p_K = \dfrac{h}{2i\pi}\cdot\dfrac{\partial}{\partial q_K}$ precisely as one does with ordinary numbers, except that one must observe the non-commutativity of the order of multiplication that corresponds to (4) (1). According to (4), every q_K commutes with every q_L, but every q_K will commute with every *other* p_L only when $L \neq K$. In general, we would like to call two functions $q'_K(p,q)$ and $p'_K(p,q)$ *canonically conjugate* when the substitution of $p_K = \dfrac{h}{2i\pi}\cdot\dfrac{\partial}{\partial q_K}$ produced the operator relations:

$$p'_K\left(q,\frac{h}{2i\pi}\frac{\partial}{\partial q_K}\right)\cdot q'_K\left(q,\frac{h}{2i\pi}\frac{\partial}{\partial q_K}\right)\psi - q'_K\left(q,\frac{h}{2i\pi}\frac{\partial}{\partial q_K}\right)\cdot p'_K\left(q,\frac{h}{2i\pi}\frac{\partial}{\partial q_K}\right)\psi = \frac{h}{2i\pi}\psi\,,$$

etc., or when written more concisely:

$$p'_K\,q'_K - q'_K\,p'_K = \frac{h}{2i\pi}\,, \qquad \text{etc.,} \quad \text{as in (4).} \tag{4$'$}$$

---

(1) The operator calculus for "q-numbers" was introduced by **P. A. M. Dirac** as the formal foundation for quantum theory. Proc. Roy. Soc. **109** (1925), pp. 642; *ibid.* **100** (1926), pp. 561. See, moreover, **C. Eckhart**, "Operator-calculus," Phys. Rev. **28** (1926), pp. 711.

As an example, we consider a system $2N$ quantities q_K and p_K that we assume to be canonically conjugate [i.e., ones that are subject to eq. (4)] and construct the following quantities from them:

$$q'_K = \sqrt{q_K}\, e^{-2i\pi p_K/h}, \qquad p'_K = \frac{2i\pi}{h} \cdot e^{+2i\pi p_K/h} \sqrt{q_K}\,, \tag{5}$$

which we assert are canonically conjugate i.e., subject to eq. (4′). In order to prove that we first consider the equation:

$$\left.\begin{aligned} e^{\pm 2i\pi p_K/h} \cdot f(q_1,\dots,q_N) &= e^{\pm\frac{\partial}{\partial q_K}} f(q_1,\dots,q_N) \\ &= \left\{1 \pm \frac{1}{1!}\frac{\partial}{\partial q_K} + \frac{1}{2!}\frac{\partial^2}{\partial q_K^2} \pm \cdots,\, f(q_1,\dots,q_K,\dots,q_N)\right\} \\ &= f(q_1,\dots,q_K \pm 1,\dots,q_N), \end{aligned}\right\} \tag{6}$$

in which the last one follows from **Taylor**’s theorem. We will then have:

$$\begin{aligned} \{p'_K\, q'_K - q'_K\, p'_K, \psi\} &= \left\{\frac{2i\pi}{h} e^{+2i\pi p_K/h} \sqrt{q_K}\sqrt{q_K}\; e^{-2i\pi p_K/h}\, \psi\right\} - \left\{\frac{2i\pi}{h} \sqrt{q_K}\, e^{-2i\pi p_K/h}\; e^{+2i\pi p_K/h} \sqrt{q_K}\, \psi\right\} \\ &= \frac{2i\pi}{h} e^{-2i\pi p_K/h}\, q_K\, e^{-2i\pi p_K/h}\, \psi(q_1,\dots,q_K,\dots,q_N) - \frac{2i\pi}{h} q_K\, \psi(q_1,\dots,q_K,\dots,q_N) \\ &= \frac{2i\pi}{h} e^{2i\pi p_K/h}\, q_K\, \psi(q_1,\dots,q_K - 1,\dots,q_N) - \frac{2i\pi}{h} q_K\, \psi(q_1,\dots,q_K,\dots,q_N) \\ &= \frac{2i\pi}{h} (q_K + 1)\, \psi(q_1,\dots,q_K + 1 - 1,\dots,q_N) - \frac{2i\pi}{h} q_K\, \psi(q_1,\dots,q_K,\dots,q_N) \\ &= \frac{2i\pi}{h}\, \psi(q_1,\dots,q_K,\dots,q_N)\ , \end{aligned}$$

so, in operator notation:

$$p'_K\, q'_K - q'_K\, p'_K = \frac{h}{2i\pi}\,.$$

One proves the remaining commutation relations (4′) for the q'_K and p'_K correspondingly. The operators q'_K and p'_K that are defined by (5) are therefore conjugate again. Conversely, one can also use the assumption that q'_K and p'_K are canonically conjugate (4) to prove that the same thing is true of p_K and q_K .

One can derive the following formulas as counterparts to (6). If τ_1, τ_2, … are any fixed quantities of dimensions p_1, p_2, …, resp., then one will have:

$$\frac{h}{2i\pi}\frac{\partial}{\partial q_K}e^{2i\pi\tau_K q_K/h}\psi(q) = \frac{h}{2i\pi}e^{2i\pi\tau_K q_K/h}\frac{\partial}{\partial q_K}\psi + \psi\,\tau_K\,e^{2i\pi\tau_K q_K/h} = e^{2i\pi\tau_K q_K/h}\left\{\frac{h}{2i\pi}\frac{\partial}{\partial q_K} + \tau_K, \psi\right\},$$

or when written as an operator equation:

$$p_K \cdot e^{2i\pi\tau_K q_K/h} = e^{2i\pi\tau_K q_K/h}\,(p_K + \tau_K)\ . \tag{7}$$

One easily proves the following, more general, equations with the help of that relation when $f(q, p)$ is any *function that can be developed in powers of* p_K :

$$f(q_1, p_1)\,e^{2i\pi(\tau_1 q_1 + \cdots + \tau_N q_N)/h} = e^{2i\pi(\tau_1 q_1 + \cdots + \tau_N q_N)/h} \cdot f(q_1, p_1 + \tau_1) \tag{8}$$

or when read in the opposite direction:

$$e^{2i\pi(\tau_1 q_1 + \cdots + \tau_N q_N)/h} f(q_K, p_K) = f(q_K, p_K - \tau_K)\,e^{2i\pi(\tau_1 q_1 + \cdots + \tau_N q_N)/h}\ . \tag{8$'$}$$

Let it be pointed out that identities that follow from the commutation relations (4) here under the reinterpretation $p_K \to \frac{h}{2i\pi}\cdot\frac{\partial}{\partial q_K}$ had already been proposed as the *fundamental law of "quantum algebra"* by **Heisenberg**, **Born**, and **Jordan**, as well as **Dirac**, before **Schrödinger** had proposed wave mechanics, in order to calculate with the help of observable quantities by using certain rules that were obtained from a bold generalization of the correspondence between classical mechanics and the quantum theory of discrete electron orbits. The operator interpretation of those quantum-algebraic relations had already been stressed by **Born** and **Wiener**, and **Lanczos** had previously introduces a continuous spatial function that was applied to the operators in a manner that did not deviate too much from **Schrödinger**. **Schrödinger** himself and **Eckart** subsequently found the complete equivalence of the "quantum mechanics" that was based upon quantum algebra and the undulatory mechanics.

In this Chapter V, we will employ the rules of calculation extensively above in order to abbreviate many otherwise-complicated formulas. The possibility of concerning ourselves with the operators that act upon the ψ-function more than that function itself is based upon just the aforementioned irrelevance of going into non-invariant functions more precisely when one would like to go beyond observable (viz., invariant) physical quantities and matters of legitimacy. One will find a related overview of the quantum algebra of invariant quantities in Chap. VII.

45. Occupation numbers as coordinates. – In what follows, we will treat the *transition probability* from a state *n* to another state *m* in a mechanical system, say, a molecule that consists of one or more atoms or a solid or gaseous aggregate that consists of many molecules, etc. To that end, it would be useful for us to take the system, which we would like to simply call *the atom*, to be a member of a large ensemble of atoms of the same type (i.e., they possess the same Hamiltonian and are mutually independent) in which at any time, N_1 atoms lie in state 1, N_2 atoms are state 2, etc. $N_1 + N_2 + \ldots = N$ is the total number of atoms in the ensemble. We further consider a second, third, etc., ensemble of N such atoms, but for which occupancy of the states is N_k', N_k'', …, resp. Let the probability of finding an ensemble in that set of all ensembles being in the occupancy state N_k' at time t be called $|\psi(N_k', t)|^2$, and let the probability of an ensemble of the set being in the state N_k'' at that same time be $|\psi(N_k'', t)|^2$. $\psi(N_k', t)$ and $\psi(N_k'', t)$ are the associated probability amplitudes. If the value of $|\psi(N_k', t)|^2$ decreases up to time t_2, while that of $|\psi(N_k'', t)|^2$ increases, then that means that there will be a *transition probability* from the occupancy state N_k' to the occupancy state N_k''.

Whereas up to now ψ was *always* considered to be a function of the coordinates q and t, like **Dirac** ([1]), we have arrived at probability amplitudes ψ in which the *occupation numbers* N_k appear as coordinates. We will arrive at their connection with Chapter IV (especially no. **41**), in which ψ was generally represented in the form:

$$\psi(q, t) = \sum_k b_k(t)\psi_k(q)$$

as a solution of the fundamental equation ([2]):

$$\frac{h}{2i\pi}\dot{\psi} + \{H, \psi\} = 0\,,$$

when we interpret the magnitude $|b_k(t)|^2$ as the number of atoms that are found in precisely the state k, such that:

$$|b_k(t)|^2 = N_k(t) \qquad \text{and} \qquad \sum_k |b_k(t)|^2 = \sum_k N_k(t) = N = \text{const.}$$

Since $b_k(t)$ itself can still be complex, one sets:

$$b_k = \sqrt{N_k}\, e^{-2i\pi\Theta_k/h}, \quad \text{with} \quad \sqrt{N_k} = \text{amplitude}, \quad \Theta_k = \text{phase}, \tag{9}$$

and we will get:

$$\psi(b_1, b_2, \ldots, q, t) = \sum_k b_k\psi_k\,, \dot{\psi} = \sum_k \dot{b}_k\psi_k\,.$$

---

(1) **P. A. M. Dirac**, Proc. Roy. Soc. **114** (1927), pp. 243.

(2) The Fourier coefficients $b_k(t)$ are rapidly-varying, $b_k = a_k \exp(-2i\pi E_k t/h)$, as opposed to the slowly-varying $a_k(t)$. The $\psi_k(t)$ are the eigen-solutions to the conservative equation $\{H^0 - E, \psi\} = 0$, in which H^0 is the "unperturbed" part of H, and $F = H - H^0$ is the perturbing function (see number **41**).

Due to the equation for the time variation of the *b* that was derived in no. **41**:

$$\frac{h}{2i\pi}\dot{b}_k + \sum_n H_{kn}\,b_n = 0\,, \tag{9′}$$

that can also be written in the form:

$$\psi = \sum_k b_k\,\psi_k\,, \qquad \frac{h}{2i\pi}\dot{\psi} + \sum_k\sum_n \psi_k H_{kn}\,b_n = 0\,. \tag{10}$$

In that way, we would like to split $H(q, t, p)$ into $H^0(q, p) + F(q, t)$, in such a way that the matrix elements F_{nn} will vanish, since the part of the perturbation energy $F(q, t)$ that would give rise to non-vanishing F_{nn} is included in H^0. One will then have:

$$H_{kn} = F_{kn} \quad \text{for} \quad k \neq n\,, \qquad H_{nn} = E_n \quad \text{and} \quad F_{nn} = 0\,. \tag{10′}$$

Now, $\frac{h}{2i\pi}\cdot b_k^*$ is defined to be the impulse conjugate to the coordinate b_k, i.e., $\frac{h}{2i\pi}\cdot b^*$ shall be an abbreviation for the operator $\frac{h}{2i\pi}\cdot \partial_k$, and b^* shall mean an abbreviation for $\partial / \partial b_k$. One will then have the operator identities [cf., (2), (2′), and (4)]:

$$\left.\begin{array}{ll} b_k^*\,b_k - b_k\,b_k^* = 1, & b_k^*\,b_k - b_k\,b_k^* = 0 \quad (l \neq k), \\ b_k\,b_l - b_l\,b_k = 0\,, & b_k^*\,b_l^* - b_l^*\,b_k^* = 0\,, \end{array}\right\} \tag{11}$$

for

$$b_k^* = \frac{\partial}{\partial b_k}\,.$$

We now understand $H(b, b^*)$ to mean the operator:

$$H(b, b^*) = \sum_k\sum_n b_k\,H_{kn}\,b_n^*\,, \tag{12}$$

and when that is applied to $\psi = \sum_l b_l\,\psi_l$, since applying $b_n^* = \partial / \partial b_k$ to $\psi_l(q)$ will give zero, we will find that:

$$\{H, \psi\} = \sum_k\sum_n b_k\,H_{kn}\,\psi_n\,,$$

such that (10) will go take on the form:

$$\left\{\frac{h}{2i\pi}\frac{\partial}{\partial t}+H,\psi\right\}=0 \qquad \text{for} \qquad \psi(b,q)=\sum_n b_n\psi_n(q)\,. \tag{13}$$

Formally, that equation has the appearance as **Schrödinger**'s fundamental equation, except that in the latter, ψ is a function of q and t, while in (13) it is a function of q and b_1, b_2, …, with time entering implicitly in the $b_k(t)$. In what follows, we will be less concerned with the dependency on q than we will be on the dependency on the b. Since the operator $H(b, b^*)$ in the differential equation (13) does not include the coordinate q, the solutions ψ of (13) will be simply functions $\psi(b_1, b_2, \ldots)$ in which each b contains time implicitly.

Now, **Dirac** further went from the canonical variables b_k and $\frac{h}{2i\pi}\cdot b_k^* = \frac{h}{2i\pi}\cdot\frac{\partial}{\partial b_k}$ to new canonical coordinates and impulses (operators), and indeed in such a way that $N_k = |\,b_k\,|^2$ would now take on the role of coordinates. That was achieved by the transformation:

$$b_k = \sqrt{N_k}\,e^{-2i\pi\Theta_k/h}, \qquad b_k^* = e^{2i\pi\Theta_k/h}\sqrt{N_k} \tag{14}$$

to the new canonical variables N_k and $\Theta_k = \frac{h}{2i\pi}\cdot\frac{\partial}{\partial N_k}$. The fact that the commutation rules:

$$\left.\begin{array}{ll}\Theta_k N_k - N_k\Theta_k = \dfrac{h}{2i\pi}, & \Theta_k N_l - N_l\Theta_k = 0 \quad \text{for} \quad l \neq k \\ N_k N_l - N_l N_k = 0, & \Theta_k\Theta_l - \Theta_l\Theta_k = 0\end{array}\right\} \tag{14′}$$

do, in fact, arise from the commutation rules (1) for the b_k under the transformation that was given (11) follows in the same way that (4′) followed from (5).

In the new canonical variables N_k Θ_k, the operator $H(b, b^*)$ in (12) will now become the operator:

$$H(N,\Theta) = \sum_k\sum_n \sqrt{N_k}\,e^{-2i\pi\Theta_k/h}\cdot H_{kn}\,e^{2i\pi\Theta_k/h}\sqrt{N_n} \qquad \text{with} \qquad \Theta_k = \frac{h}{2i\pi}\frac{\partial}{\partial N_k},$$

which will make the operator in the fundamental (13) go to:

$$\frac{h}{2i\pi}\frac{\partial}{\partial t}+\sum_k\sum_n H_{kn}\sqrt{N_k}\,e^{-2i\pi\Theta_k/h}\,e^{2i\pi\Theta_k/h}\sqrt{N_n} = 0\,. \tag{14″}$$

If one applies it to a function Ψ then Ψ will be a function of the *continuous* coordinates $N_1, N_2, \ldots,$ each of which can take on all values between 0 and ∞. The restriction of the coordinate values N_k to *whole numbers*, or rather, restricting the changes in the N_k to whole numbers, is just a result of quantum mechanics that can be derived immediately.

46. Transition probabilities. – The last equation can be put into the form:

$$\frac{h}{2i\pi}\dot{\Psi}(N_1,\ldots,N_k,\ldots,N_n,\ldots)+\sum_k\sum_n H_{kn}\sqrt{N_k}\cdot\sqrt{N_k+1}\cdot\Psi(N_1,\ldots,N_k-1,\ldots,N_n+1,\ldots) = 0 \quad (15)$$

with the help of formulas (6). Only the terms with $k \neq n$ are written in out the double sum. For $k = n$, one must add the simple sum:

$$\sum_k H_{kk}\sqrt{N_k}\sqrt{N_k}\cdot\Psi(N_1,\ldots,N_k,\ldots) = \sum_k E_k\, N_k \Psi(N_1,\ldots,N_k,\ldots). \quad (15')$$

We regard Ψ ($N_1, N_2, \ldots, N_k, \ldots$) (see also Example 3 in no. **40**) as the probability amplitude that an atomic ensemble that belongs to a set of atomic ensembles will be found in the occupation state $N_1, N_2, \ldots$ at time t. Ψ ($N_1, \ldots, N_k - 1, \ldots, N_n + 1, \ldots$ is then the probability amplitude at that same time for an ensemble in the set to have an occupation state that differs from the first one by the fact that there is one less atom in the state k and one more atom in the state n. The *temporal variation* of the former probability amplitude is coupled with the amplitudes of the latter probabilities by equation (15). In order to simplify the notation, we now understand Ψ_N and Ψ_M to means the function Ψ_N ($N_1, N_2, \ldots$) and Ψ_M ($M_1, M_2, \ldots$), in place of (15), one can write, more concisely:

$$0 = \frac{h}{2i\pi}\dot{\Psi}_N + \sum_M \mathfrak{H}_{NM}\cdot\Psi_M \ , \quad (16)$$

in which the sum extends over all distributions M that deviate from the distribution N by the fact that one N is increased by 1 and another is reduced by 1, and in that way, there will be a term $N = M$ that corresponds to (15′). The meaning of the coefficients $\mathfrak{H}_{NM}$ can be read off by a comparison with (15), (15′). For example, from (15′), one will have $\mathfrak{H}_{NN} = \sum_k N_k\, E_k$, which we would like to also write as E_n :

$$\mathfrak{H}_{NN} = E_N = \sum_k N_k\, E_k \quad (16')$$

as the total energy of the atom in the occupancy state N.

One can go from the equation (16) for rapidly-varying functions Ψ_N to the one for slowly-varying functions Φ_N when one introduces the Ansatz:

$$\Psi_N = \Phi_N \cdot e^{2i\pi E_N t/h}. \tag{17}$$

One will then get [cf., the transition from (28) to (28′) in no. **41**] the equation for Φ :

$$0 = \frac{h}{2i\pi}\dot{\Phi}_N + \sum_M \mathfrak{F}_{NM}\Phi_M\, e^{2i\pi(E_N - E_M)t/h} \qquad \text{with} \quad \mathfrak{F}_{NM} = \begin{cases} \mathfrak{H}_{NM} & \text{for} \quad N \neq M \\ 0 & \phantom{\text{for}} \quad N = M. \end{cases} \tag{18}$$

The *physical meaning* of the $\mathfrak{H}_{NM}$ ($\mathfrak{F}_{NM}$, resp.) also corresponds to the meaning of the H_{nm} (F_{nm}, resp.) that was treated in no. **41**. The integration of (18) for the time interval 0 to t [cf., the integral of (28′) in no. **41**] will lead to the series:

$$\Phi_N(t) - \Phi_N(0) = \sum_M f_{NM}(t)\,\Phi_M(0)\,, \tag{19}$$

and in the first approximation, namely, for small time intervals over which the Φ_M vary only slightly:

$$f_{NM}^{(1)}(t) = -\frac{2i\pi}{h}\int_0^t \mathfrak{F}_{NM}\, e^{2i\pi(E_N - E_M)t/h}\, dt\,. \tag{20′}$$

If the time interval from 0 to t is taken to be small enough that $\mathfrak{F}_{NM}$ can also be regarded as constants then one will get:

$$f_{NM}^{(1)}(t) = \mathfrak{F}_{NM}\cdot\left(1 - e^{2i\pi(E_N - E_M)t/h}\right) : (E_N - E_M) \tag{21}$$

upon integration. The coefficients of the development $f_{NM}(t)$ in the series (19) are the *probability amplitudes,* while the $|f_{NM}|^2 = w_{NM}$ are the *probabilities* that an atomic ensemble will jump from the occupation state M to N during the time t. Precisely as in no. **41**, there is a *law of interference of probabilities* that follows as a consequence of the law of addition (19) for probability amplitudes.

The peculiarity of the problem that was treated here of the transition of an atomic ensemble from an occupancy state $M = (M_1, M_2, \ldots)$ to the occupancy state $N = (N_1, N_2, \ldots)$ now implies that the quantities $\mathfrak{H}_{NM}$ ($\mathfrak{F}_{NM}$, resp.) that lead to the transition amplitudes f_{NM} will differ from zero only the following case, which one reads off from (15):

$$\left.\begin{array}{c|c|c} \text{Final distribution } N & \text{Initial distribution } M & \mathfrak{F}_{NM} = \mathfrak{H}_{NM} \\ = N_1, \ldots, N_k, \ldots, N_n, \ldots & = N_1, \ldots, N_k - 1, \ldots, N_n + 1, \ldots & = F_{kn}\sqrt{N_k}\sqrt{N_n + 1} \end{array}\right\} \tag{22}$$

i.e., the coordinates N_k, which were continuous from the outset, can either remain the same or change by ± 1, so in succession they can jump by only *integer multiples* of 1. In detail, (22) means the following: Under the transition from initial distribution of the atomic ensemble to the final distribution, *one* atom will jump from the state n to the state k. The asymmetry in the transition producer $F_{kn}\sqrt{N_k}\sqrt{N_k+1}$ is only apparent, since one can imagine that the atom that is making the jump is one of the (N_n + 1) atoms that were found in the state n to begin with, and that the same atom will be one of the N_k atoms that are found in the state k at the end. (22) is also equivalent to the following table:

$$\left.\begin{array}{c|c|c} \text{Final distribution } N & \text{Initial distribution } M & \mathfrak{F}_{NM} = \mathfrak{H}_{NM} \\ = N_1, \ldots, N_k + 1, \ldots, N_n - 1, \ldots & = N_1, \ldots, N_k, \ldots, N_n, \ldots & = F_{kn}\sqrt{N_k+1}\sqrt{N_n} \end{array}\right\} \qquad (22')$$

For example, if all $N_k = 0$, except $N_n = 1$, to begin with then then the transition probability for the atom to go from the state n to the state k is calculated from $\mathfrak{F}_{NM} = F_{kn} \cdot \sqrt{1}\sqrt{1} = F_{kn}$, i.e., simply from the matrix element of the perturbation energy. One sees quite clearly here that in order for the transition from n to k to be possible, it is not necessary for, say, one or more atoms to lie in the state k *before* the transition (as is sometimes incorrectly concluded from wave mechanics).

47. Interaction of light and electrons. – In conjunction with **P. Dirac** (1), we will now specialize the perturbing potential F (q, t) and investigate the interaction of light quanta and electrons in order to arrive at a quantum theory of *absorption, emission, scattering of light,* and *dispersion.* In so doing, the simplest formulas from the relativistic classical mechanics of electrons in an electromagnetic field might be assumed. They will be recalled in Chapter VI, moreover.

In a perturbing field with a scalar potential φ and a vector potential with components $\mathfrak{A}_x$, $\mathfrak{A}_y$, $\mathfrak{A}_z$, the energy equation reads:

$$\left(p_x + \frac{\varepsilon}{c}\mathfrak{A}_x\right)^2 + \cdots - \frac{1}{c^2}(E + \varepsilon\varphi)^2 + \mu^2 c^2 = 0 \, .$$

Therefore, the Hamiltonian function is now:

$$H = E = c \cdot \left[\mu_0^2 c^2 + \left(p_x + \frac{\varepsilon}{c}\mathfrak{A}_x\right)^2 + \cdots\right]^{1/2} - \varepsilon\varphi \; .$$

We consider the *unperturbed* case to be the presence of just the scalar potential φ, which shall be time-independent, moreover (e.g., motion in the field of a nucleus) with the Hamiltonian function:

$$H_0 = E_0 = c\sqrt{\mu_0^2 c^2 + p^2} - \varepsilon\varphi \, . \qquad (23)$$

---

(1) **P. A. M. Dirac**, Proc. Roy. Soc. **114** (1927), pp. 243.

When one neglects v^2 / c^2 in comparison to 1 and the terms in $\mathfrak{A}^3$ then one will get:

$$H = E = H_0 + \frac{\varepsilon}{c}(\mathfrak{v}\,\mathfrak{A}) + \frac{\varepsilon^2}{2\mu c^2}\mathfrak{A}^2 \tag{24}$$

approximately.

That equation represents a mixing of the corpuscular theory of matter with the field theory of light. In order to arrive at a consistent field theory of light + matter, it is next necessary to describe the interaction of electrons and light corpuscles (light quanta) by a point-mechanical **Hamiltonian** that is then translated into wave mechanical form. In that way, the atom, with its degrees of freedom, is contrasted with the light field, with its degrees of freedom, which is introduced mathematically as a system of harmonic oscillators that interact perturbatively with the atom, in a certain sense. Physically, those oscillators are the elementary *eigen-oscillations* of the light that fills up the space. For our purposes, it is especially appropriate to use **Laue**'s elementary bundle of rays (cf., no. **17**, and further no. **9**), such that there will be:

$$\frac{8\pi\nu^2}{c^3}V\,d\nu \tag{25}$$

elementary ray bundles in a volume V and in the frequency interval $d\nu$, which agrees with the **Jeans** number for the eigenfrequencies of *both* directions of polarization. That number of elementary ray bundles with *one* direction of polarization that then falls upon an aperture of solid angle $d\omega$ (instead of 4π) will then be:

$$\sigma = \frac{\nu^2}{c^3}V\,d\nu\,d\omega \qquad \text{elementary ray bundles in } V\,d\nu\,d\omega\,. \tag{25′}$$

In the corpuscular theory of light, each of them will be occupied by a while number N_k of energy quanta $h\nu_k$. However, we *would not* like to assume that N_k is a whole number here, but initially assume that N_k lies in the continuous domain from 0 to ∞. The state of light in space will then be characterized by all (possibly rational) numerical values N_k of light units $h\nu_k$ which lie on the individual (on the k^{th}) ray bundles and the phases Θ_k of those ray bundles.

Now, in order to obtain the perturbation between the latter and the atom, we next consider only the k^{th} ray bundle with the (continuously-varying) occupation number N_k and phase Θ_k , and ask what vector potential $\mathfrak{A}^k$ that it would generate might be. The mean energy density of the ray bundle that is generated in the volume V is $N_k\,h\,\nu_k / V$. On the other hand, the mean electromagnetic field density of the ray with frequency ν is equal to:

$$\frac{1}{8\pi}\left(\overline{\mathfrak{E}^2} + \overline{\mathfrak{H}^2}\right) = \frac{1}{8\pi}\overline{\mathfrak{E}^2} = \frac{1}{4\pi c^2}\left(\overline{\dot{\mathfrak{A}}}\right)^2 = \frac{1}{4\pi c^2}\frac{4\pi^2\nu^2}{c^2}\overline{\mathfrak{A}^2} = \frac{\pi\nu^2}{c^2}\overline{\mathfrak{A}^2}\,.$$

Upon setting both densities equal to each other, it will follow that $\overline{\mathfrak{A}_k^2} = \frac{c^2 h}{\pi \nu} \cdot \frac{N_k}{V}$, and when we make $\mathfrak{A}^k$ periodic with the phase factor cos $(2\pi \Theta_k / h)$, we will finally have:

$$|\mathfrak{A}^k| = \sqrt{N_k} \cos\left(\frac{2\pi\,\Theta_k}{h}\right) \cdot 2 \sqrt{\frac{h \cdot c^2}{2\pi \cdot \nu_k V}} \,. \tag{26}$$

$\mathfrak{A}^k$ itself points in the direction of the electric vector $\mathfrak{E}^k = - \dot{\mathfrak{A}}^k / c$. The total vector potential $\mathfrak{A}$ is the superposition of all contributions k. In that way, $\mathfrak{A}$ is expressed in terms of the corpuscular data N_k and Θ_k.

Before substituting that in the Hamiltonian (24), we need to symmetrize:

$$2 \cdot \sqrt{N_k} \cos\left(\frac{2\pi\,\Theta_k}{h}\right) = \sqrt{N_k}\, e^{-2i\pi\Theta_k/h} + e^{2i\pi\Theta_k/h} \sqrt{N_k} = b_k + b_k^* \qquad \text{[cf., (14)].}$$

With the abbreviations:

$$|\mathfrak{v}| \cdot \cos(\mathfrak{A}^k, \mathfrak{v}) = \mathfrak{v}_k \qquad \text{and} \qquad \cos(\mathfrak{A}^k, \mathfrak{A}^l) = \cos \alpha_{kl} \,,$$

the Hamiltonian (24) will then become:

$$H = H_0 + \frac{\varepsilon}{c} \sum_k \sqrt{\frac{h c^2}{2\pi V \nu_k}} \cdot \mathfrak{v}_k \cdot (b_k + b_k^*) + \frac{\varepsilon^2}{2\mu c^2} \sum_k \sum \frac{h c^2}{2\pi \sqrt{\nu_k \nu_l}} \cdot (b_k + b_k^*) \cdot (b_l + b_l^*) + \sum_k N_k h \nu_k \,, \tag{27}$$

when we add the unperturbed light energy.

With **Dirac**, one can understand that equation as follows: It is the Hamiltonian for a mechanical system with infinitely-many degrees of freedom, namely, the cavity V with its eigenfrequencies (ray bundles k = 1, 2, 3, …, ∞) in a "configuration" in which the eigenfrequency coordinates possess the values b_1, b_2, … and the canonically-conjugate impulses possess the values:

$$\frac{h}{2i\pi} b_1^* , \qquad \frac{h}{2i\pi} b_2^* , \qquad \ldots,$$

or also when one employs the transformation:

$$b_k = \sqrt{N_k}\, e^{-2i\pi\Theta_k/h} , \quad b_k^* = e^{2i\pi\Theta_k/h} \sqrt{N_k} \,, \tag{28}$$

for the values of the coordinates N_k and conjugate impulses Θ_k . In that way, one will arrive at the coordinates and impulses of the electron that are included in H_0 and the factors $\mathfrak{v}_k$.

48. Transition probabilities (^{1}). – The transition from the point-mechanical Hamiltonian (27) to the wave-mechanical operator comes about by reinterpreting the impulses b_k^* in terms of the operators $\partial / \partial b_k$ [the impulses Θ_k in terms of $\frac{h}{2i\pi}\frac{\partial}{\partial N_k}$, resp.] and constructing the wave equation (16):

$$-\frac{h}{2i\pi}\dot{\Psi}_N = \sum_M \mathfrak{H}_{NM}\cdot\Psi_M\,, \tag{28′}$$

in which Ψ is a function of the continuous coordinates b_k or N_k . When we observe the conversion (6):

$$e^{\pm 2i\pi\Theta_k/h} f(N_1,\ldots,N_k,\ldots) = f(N_1\,,\,\ldots,\,N_k \pm 1,\,\ldots)$$

and write simply Ψ, instead of Ψ (N_1 , N_2 , …) and Ψ (N_k + 1), instead of Ψ (N_1 , N_2 , …, N_k + 1, …), for brevity, then the equation above can be written out in detail as:

$$\left.\begin{aligned} -\frac{h}{2i\pi}\dot{\Psi} = \; & H_0\,\Psi + \sum_k N_k\,h\nu_k\Psi \\ & + \varepsilon\cdot\sum_k\sqrt{\frac{h}{2\pi V\,\nu_k}}\,\mathfrak{v}_k\cdot\Big[\sqrt{N_k}\,\Psi(N_k-1)+\sqrt{N_k+1}\,\Psi(N_k+1)\Big] \\ & + \frac{\varepsilon^2}{2\mu}\sum_k\sum_l\frac{h}{2\pi V}\frac{\cos\alpha_{kl}}{\sqrt{\nu_k\,\nu_l}} \\ & \Big[\sqrt{N_kN_l}\cdot\Psi(N_k-1,N_l-1)+\sqrt{N_k(N_l+1)}\cdot\Psi(N_k-1,N_l+1) \\ & +\sqrt{(N_k+1)N_l}\cdot\Psi(N_k+1,N_l-1)+\sqrt{(N_k+1)(N_l+1)}\cdot\Psi(N_k+1,N_l+1)\Big]. \end{aligned}\right\} \tag{29}$$

We already see here that the numbers N_k always change by only ± 1, which corresponds with the considerations in no. **46**, i.e., that only *whole light quanta* can ever be exchanged. From the result that was obtained, with **Dirac**, we will read off the probabilities of the various optical processes of emission, absorption, scattering, etc. However, in order to do that we must make an extension in scope. Namely, operators H_0 and $\mathfrak{v}_k$ that act upon Ψ in (29) include the coordinates q and impulses p of the *electron,* such that all functions Ψ that occur will include the argument q, in addition to the N_k . In no. **45**, it was shown that one could also introduce numbers N_1^ε , N_2^ε , … as

---

(^{1}) **P. A. M. Dirac**, Proc. Roy. Soc. **114** (1927), pp. 243.

coordinates of the electron, instead of q, which refer to the distribution over the possible quantum states 1, 2, … for an entire *ensemble* of bound electrons of the same type. Correspondingly, Ψ will be a continuous function of the arguments:

$$\Psi(N_1^\varepsilon, N_2^\varepsilon, \ldots, N_1, N_2, \ldots) \tag{30}$$

for a family of electrons and light quanta. If we restrict ourselves to the case in which *only one electron is present, and indeed in the initial state m and in the initial state n* then the operator H on the right-hand side of (29), which depends upon q and p must be replaced with its *matrix element* $\mathfrak{H}^{nm}$ (see pp. 92, footnote), so we *must now write* $\mathfrak{v}_k^{nm}$ *for* $\mathfrak{v}_k$ and $(H_0)^{nm} = 0$ *for* H_0 . (28) will then assume the following form:

$$-\frac{h}{2i\pi}\dot{\Psi}\begin{pmatrix} n \\ N \end{pmatrix} = \sum_M \mathfrak{H}_{NM}^{nm} \cdot \Psi\begin{pmatrix} m \\ M \end{pmatrix}, \tag{31}$$

in which the upper index n (m, resp.) shall refer to the electron state, while the lower one N (M, resp.) shall refer to the light coordinates. The values $\mathfrak{H}_{NM}$ are read off from (29), and $\mathfrak{H}_{NM}^{nm}$ shall give their *matrix elements* relative to the eigenfunctions ψ_n (q) and ψ_m (q) of the electron whose eigenvalues are equal to E_n and E_m . We then have taken an equation for the rapidly-varying quantities Ψ that is analogous to equation (16). The Ansatz [cf., (17)]:

$$\Psi\begin{pmatrix} n \\ N \end{pmatrix} = \Phi\begin{pmatrix} n \\ N \end{pmatrix} e^{2i\pi t(E_N + E_n)/h}$$

then leads to the equation for the slowly-varying quantities Φ [cf., (18)]:

$$-\frac{h}{2i\pi}\dot{\Phi}\begin{pmatrix} n \\ N \end{pmatrix} = \sum_M \mathfrak{F}_{NM}^{nm} \cdot \Phi\begin{pmatrix} m \\ M \end{pmatrix} e^{2i\pi t(E_N - E_M + E_n - E_m)/h}, \tag{32}$$

in which only the terms with $N \neq M$ are of any interest to us since they are definitive of the light conversions.

The integration over time then leads to a series [cf., (19)]:

$$\Phi\begin{pmatrix} n \\ N \end{pmatrix}_{t=0} - \Phi\begin{pmatrix} n \\ N \end{pmatrix}_t = \sum_M f_{NM}^{nm}(t) \cdot \Phi\begin{pmatrix} m \\ M \end{pmatrix}_{t=0}, \tag{33}$$

whose coefficients $f_{NM}^{nm}(t)$ mean the *transition amplitudes* from the light state M to N with a simultaneous jump of the electron from m to n. The magnitudes of their squares are the associated *transition probabilities* $w_{NM}^{nm}(t)$. In detail, upon considering the various sums of terms in (29) in the first approximation, one will get the emission, absorption, and scattering probabilities, and in the second approximation, one will get the dispersion phenomena, as will be shown in what follows in connection with **Dirac.**

49. Emission, absorption, scattering (1). – The integration of (32) leads to the transition amplitudes [cf., (21)]:

$$f_{NM}^{nm}(t) = \mathfrak{F}_{NM}^{nm}\cdot\left(1-e^{2i\pi(E_N+E_n-E_M-E_m)t/h}\right):(E_N+E_n-E_M-E_m)$$

in the *first approximation* and the transition probabilities:

$$w_{NM}^{nm}(t) = |F_{NM}^{nm}|^2\cdot 2\left[1-\cos\frac{2\pi t}{h}(E_N+E_n-E_M-E_m)\right]:(E_N+E_n-E_M-E_m)^2 . \tag{34}$$

The *emission* of a light quantum $h\nu_k$ under an atomic jump $m\to n$ corresponds the radiation from the final state $N = N_1, N_2, \ldots$ to the initial state $M = N_1, \ldots, N_k - 1, \ldots$ The associated coefficient $\mathfrak{F}_{NM}^{nm} = \mathfrak{H}_{NM}^{nm}$ is, as one can read off from (29):

$$\mathfrak{H}_{NM}^{nm} = \varepsilon\sqrt{\frac{h}{2\pi V\nu_k}}\,\mathfrak{v}_k^{nm}\sqrt{N_k} \qquad \begin{pmatrix} E_M=(N_k-1)\,h\nu_k \\ E_N=N_k\,h\nu_k \end{pmatrix}.$$

With a different notation for the initial and final state, that is equivalent to:

$$\mathfrak{H}_{NM}^{nm} = \sqrt{\frac{h}{2\pi V\nu_k}}\cdot\dot{\mathfrak{M}}_k^{nm}\sqrt{N_k+1} \qquad \begin{pmatrix} E_M=N_k\,h\nu_k \\ E_N=(N_k+1)\,h\nu_k \end{pmatrix} \tag{34′}$$

when one introduces the electric moment $\mathfrak{M}=\varepsilon\,\mathfrak{r}$, $\dot{\mathfrak{M}}=\varepsilon\,\mathfrak{v}$. If one further sets:

$$E_m-E_n=\nu_{nm}\cdot h \qquad \begin{pmatrix} M\to N \\ m\to n \end{pmatrix}, \qquad E_M-E_N=-h\,\nu_k ,$$

to abbreviate, then one will arrive at the *emission probability:*

$$w_{NM}^{nm}(t) = \frac{1}{V\pi h\nu_k}|\dot{\mathfrak{M}}_k^{nm}|^2\,(N_k+1)\left[1-\cos 2\pi t\,(\nu_{nm}-\nu_k)\right]:(\nu_{nm}-\nu_k)^2 . \tag{35}$$

---

(1) **P. A. M. Dirac**, Proc. Roy. Soc. **114** (1927), pp. 243; moreover, see **A. Landé**, Zeit. Phys. **42** (1927), pp. 835; **L. Landau**, *ibidem*, **45**, pp. 430; **G. Wentzel**, *ibidem*, **43** (1927), pp. 524; **E. Fues**, *ibidem*, **43** (1927), pp. 726; **F. Bloch**, Phys. Zeit. **29** (1928), pp. 58; **J. C. Slater**, Proc. Mat. Acad. Amer. **13** (1927), pp. 7.

It is not proportional to the number N_k of light quanta in the initial state, but to $N_k + 1$. One calls the part that is proportional to N_k the *induced emission* of radiation, and the part that is proportional to 1, the *spontaneous emission*. The emitted light quantum ν_k does not need to possess, say, the same frequency as the ν_{nm} that belongs to the atomic jump, but generally the emission probability will become larger as the denominator $(\nu_{nm} - \nu_k)^2$ becomes smaller. *The* ***Bohr*** *frequency condition is not sharply fulfilled then*. (35) will then give the intensity distribution of the broad spectral line about the maximum $\nu_k = \nu_{mn}$ that is emitted by the atomic jump $m \to n$.

One will get the total spontaneous emission probability in the volume V in a solid angle $d\omega$ and with frequencies in the interval $d\nu$ from (35) when one replaces $N_k + 1$ with 1 and multiplies by the number $\sigma_k = \nu_k^2 / c^2 \cdot V\, d\nu_k\, d\omega_k$ of the associated elementary ray bundles (25′). V drops out of (35) with that:

$$d\omega_k \cdot \frac{1}{\pi h c^3} |\dot{\mathfrak{M}}_k^{nm}|^2 \cdot \nu_k \cdot \frac{1 - \cos 2\pi t\,(\nu_{nm} - \nu_k)}{(\nu_{nm} - \nu_k)^2}\, d\nu_k \;.$$

If one would like to obtain the emission probability for all colors ν_k by integrating over ν_k from 0 to ∞ then that will show that the integral does not converge, because the contribution of the higher frequencies will decrease too slowly (like $d\nu_k / \nu_k$). That will happen because for very high frequencies, i.e., small wavelengths, it is not permissible to replace the field that acts on the electron in its instantaneous position with the field that acts on it in its rest position in the point-mechanical Ansatz. If one imagines getting around the difficulty regard convergence by a suitable modification for large ν_k then only the ν_k in the neighborhood of ν_{mn} would contribute essentially to the integral, such that one can replace the *factor* ν_k with ν_{mn} in the integration. One will then find the emission probability for arbitrary colors in the $d\omega_k$-cone:

$$d\omega_k \frac{1}{\pi h c^3} |\dot{\mathfrak{M}}_k^{nm}|^2\, 2\pi^2 \nu_{nm}\, t = d\omega_k \frac{2\pi \nu_{nm}\, t}{hc^3} \left| \frac{d}{dt} \mathfrak{M}_k^{nm} \right|^2 = d\omega_k \cdot \frac{8\pi^3 \nu_{nm}^3\, t}{hc^3} \left| \mathfrak{M}_k^{nm} \right|^2, \tag{36}$$

in which one identifies $\dot{\mathfrak{M}}_k^{nm}$ with:

$$\frac{d}{dt} \mathfrak{M}_k^{nm} = \frac{2i\pi}{h} \nu_{nm}\, \mathfrak{M}_k^{nm} \;.$$

One will obtain the total spontaneous emission probability of the atom that jumps from the state m to n when one replaces $d\omega_k$ with 4π, adds a factor of 2 to account for the two directions of polarization, and replaces the mean value of the k-components with a factor of 1/3:

$$\overline{|\mathfrak{M}_k^{nm}|^2} = \tfrac{1}{3} \overline{|\mathfrak{M}^{nm}|^2} \;.$$

If one then multiplies that by $h\nu_k$ then one will get the total energy per atom and per unit time that emanates from a large ensemble of atoms of the same type:

$$\frac{dE}{dt} = \frac{64\pi^4 \nu_{nm}^4}{3c^3} \overline{|\mathfrak{M}^{nm}|^2} \tag{36′}$$

under the transition $m \to n$. That expression coincides with the one that was used in no. **43**, eq. (41) for radiation and will then justify the Ansatz that was employed there for the "transition density" ρ^{nm} as the source of classical radiation.

What is definitive of the *absorption* of a light quantum $h\nu_k$ in the final state $N = N_1, N_2, \ldots$ under the radiation to the initial state $N = N_1, \ldots, N_k + 1, \ldots$ The associated coefficient H_{NM}^{nm} in (29) is:

$$H_{NM}^{nm} = \varepsilon \sqrt{\frac{h}{2\pi V \nu_k}}\, \mathfrak{v}_k^{nm} \sqrt{N_k + 1} \qquad \begin{pmatrix} E_M = (N_k + 1) \cdot h\nu_k \\ E_N = N_k \cdot h\nu_k \end{pmatrix}.$$

That is equivalent to:

$$H_{NM}^{nm} = \sqrt{\frac{h}{2\pi V \nu_k}}\, \dot{\mathfrak{M}}_k^{nm} \sqrt{N_k} \qquad \begin{pmatrix} E_M = N_k \, h\nu_k \\ E_N = (N_k - 1)\, h\nu_k \end{pmatrix}. \tag{37}$$

That differs from (34′) by the factor $\sqrt{N_k}$, instead of $\sqrt{N_k + 1}$, so the ratio of the emission probability to the absorption probability is:

$$\frac{\text{emission probability}}{\text{absorption probability}} = \frac{N_k + 1}{N_k}. \tag{38}$$

One can employ that relation as the basis for **Einstein**'s proof of the **Planck** radiation law.

The *scattering* of a light quantum does not encompass the Compton effect, since v^2 / c^2 was neglected above. We consider the transition to the final state $N = N_1, N_2, \ldots$ from the initial state $M = N_1, \ldots, N_k + 1, \ldots, N_l - 1, \ldots$ In (29), that corresponds to:

$$H_{NM}^{nm} = 2\, \frac{\varepsilon^2 h}{2\mu \cdot 2\pi V} \frac{\cos\alpha_{kl}}{\sqrt{\nu_k \nu_l}} \sqrt{N_l} \sqrt{N_k + 1} \qquad \begin{pmatrix} E_M = (N_k + 1)\, h\nu_k + (N_l - 1)\, h\nu_l \\ E_N = N_k h\nu_k + N_l \, h\nu_l \end{pmatrix}.$$

Such a term will, in fact, occur twice in the double sum. Instead of it, we shall use:

$$H_{NM}^{nm} = \frac{\varepsilon^2 h \cos\alpha_{kl}}{2\pi \, \mu V \sqrt{\nu_k \nu_l}} \sqrt{N_l + 1} \sqrt{N_k} \qquad \begin{pmatrix} E_M = N_k \, h\nu_k + N_l \, h\nu_l \\ E_N = (N_k - 1)\, h\nu_k + (N_l + 1)\, h\nu_l \end{pmatrix} \tag{39}$$

for the transition from a light quantum $h\nu_k$ to a light quantum $h\nu_l$. According to (34), that will lead to the transition probability:

$$w_{NM}^{nm} = \frac{\varepsilon^4 \cos^2 \alpha_{kl}}{4\pi^2 \mu^2 \cdot V\nu_k \cdot V\nu_l} (N_l + 1) N_k \cdot 2 \cdot \frac{1 - \cos 2\pi t (\nu_l - \nu_k - \nu_{nm})}{(\nu_l - \nu_k - \nu_{nm})^2} . \tag{39′}$$

Upon multiplying by $\sigma_k = \nu_k^2 / c^3 V \, d\nu_k \, d\omega_k$ and $\sigma_l = \nu_l^2 / c^3 V \, d\nu_l \, d\omega_l$ [see (25′)], one will get the *transition probability* from the ray bundle $d\nu_k \, d\omega_k$ to the bundle $d\nu_l \, d\omega_l$:

$$\frac{\varepsilon^4 \cos^2 \alpha_{kl}}{2\pi^2 \mu^2 c^6} (N_l + 1) N_k \nu_k \nu_l \cdot \frac{1 - \cos 2\pi t (\nu_l - \nu_k - \nu_{nm})}{(\nu_l - \nu_k - \nu_{nm})^2} d\nu_k \, d\omega_k \, d\nu_l \, d\omega_l . \tag{39″}$$

The case of $\nu_{nm} = 0$ for the initial state $N_l = 0$ holds a special interest, i.e., scattering from an atom that remains unchanged into a bundle $d\nu_l \, d\omega_l$ that is free of rays. If one integrates over $d\nu_l$, in addition, and gets around the difficulty pertaining to convergence as on pp. 114 then one will get:

$$\frac{\varepsilon^4}{2\pi^2 \mu^2 c^6} \cos^2 \alpha_{kl} \, \nu_k N_k \cdot 2\pi^2 \nu_k \, t \cdot d\nu_k \, d\omega_k \, d\omega_l \tag{40}$$

for the probability of a scattering process from $d\nu_k \, d\omega_k$ to $d\omega_l$ to occur during the time t. The radiant energy that falls upon a unit area in unit time with a certain polarization is $\Re(\nu_k) \, d\nu_k \, d\omega_k$ then, and the radiation density will then be $\rho(\nu_k) = N_k h\nu_k \cdot 8\pi \nu_k^2 / c^3$ [cf., (25)], so:

$$\Re(\nu_k) = \frac{N_k h \nu_k^2}{c^2} . \tag{40′}$$

If one then substitutes $N_k = c^2 \Re(\nu_k) / h\nu_k^3$ in (40) and multiplies by $h\nu_k$ then one will get the *scattered energy* per electron per unit time from $d\nu_k \, d\omega_k$ to $d\omega_l$:

$$\frac{dE}{dt} = \frac{\varepsilon^4}{\mu^2 c^4} \cos^2 \alpha_{kl} \, \Re(\nu_k) \, d\nu_k \, d\omega_k \, d\omega_l \tag{41}$$

for an incident area density of radiation $\Re(\nu_k) \, d\nu_k \, d\omega_k$. The α_{kl} in that gives the angle between the electric vector of the incident and scattered rays. (41) duplicates a result that is known from the theory of classical oscillators.

50. Dispersion. – In order to do justice to the dispersion phenomena, according to **Dirac** ([1]), it will be necessary to consider the integration of (32) in the second approximation. In place of (32), we shall write, more briefly ([2]):

$$-\frac{h}{2i\pi}\,a_m(t) = \sum_n F_{mn}\,a_n(t)\cdot e^{2i\pi(E_m-E_n)t/h} \qquad (F_{mn} = 0 \text{ for } m \neq n). \qquad (42)$$

We shall use the solution in the first-approximation [cf., (21)]:

$$a_m(t) = a_m(0) + \sum_n F_{mn}\,a_n(0)\left(1-e^{2i\pi(E_m-E_n)t/h}\right):(E_m-E_n)$$

for the special case in which *all systems are in the initial state k*, $a_k(0) = 1$, $a_m(0) = 0$ for $m \neq k$.

With the use of the abbreviation:

$$\delta_{nk} = \begin{cases} 1 & \text{for} \quad n = k \\ 0 & \qquad\;\; n \neq k \end{cases}$$

one will then have:

$$a_m(t) = \delta_{mk} + \sum_n F_{mn}\left(1-e^{2i\pi(E_m-E_n)t/h}\right):(E_m-E_n)\;. \qquad (42')$$

If one substitutes that first approximation into the right-hand side of (42) then one will get:

$$-\frac{h}{2i\pi}\,a_m(t) = F_{mk}\,e^{2i\pi(E_m-E_k)t/h} + \sum_n F_{mn}\,F_{nk}\left(1-e^{2i\pi(E_m-E_k)t/h}\right)\cdot e^{2i\pi(E_m-E_n)t/h} :(E_m-E_k)$$

$$= \left(F_{mk} - \sum_n \frac{F_{mn}\,F_{nk}}{E_n-E_k}\right)e^{2i\pi(E_m-E_k)t/h} + \sum_n \frac{F_{mn}\,F_{nk}}{E_n-E_k}\,e^{2i\pi(E_m-E_n)t/h}\,.$$

Integrating that equation for $m \neq k$ will give:

$$a_m(t) = \left(F_{mk} - \sum_n \frac{F_{mn}\,F_{nk}}{E_n-E_k}\right)\cdot\frac{1-e^{2i\pi(E_m-E_n)t/h}}{E_m-E_k} + \sum_n \frac{F_{mn}\,F_{nk}}{E_n-E_k}\cdot\frac{1-e^{2i\pi(E_m-E_n)t/h}}{E_m-E_n}\,. \qquad (43)$$

In the second sum $\sum_k$, it cannot happen that a term will become infinite due to the vanishing of a denominator $E_r - E_s$, since the factor F_{rs} is in the numerator, and it will vanish whenever $r = s$. By contrast, in the first term on the right-hand side of (43), the denominator $E_m - E_k$ cannot be rendered harmless by a simultaneous vanishing of the numerator F_{mk}. For E_m that lie close to E_k, that first

---

([1]) **P. A. M. Dirac**, Proc. Roy. Soc. **114** (1927), pp. 710.

([2]) From eq. (42) to (44), *n* must be switched with *m* (rem. by the editor).

term will far exceed the second $\sum_n$ then. For those m for which E_m lies close to E_k, the $\sum_n$ can simply be dropped, and one will then get the sum:

$$\sum_m |a_m|^2 = \sum_m \left| F_{mk} - \sum_n \frac{F_{mn}\,F_{nk}}{E_n - E_k} \right|^2 \cdot \frac{2\left[1 - \cos\frac{2\pi}{h}(E_m - E_k)\,t\right]}{(E_n - E_k)^2}$$

for those m. If ΔE_m is the small distance between the energy levels that are distributed in the neighborhood of E_k in the sum then one can replace the $\sum$ with $\frac{1}{\Delta E_m}\int dE_m$ and integrate from $E_m = 0$ to ∞, since only the energies $E_m \approx E_k$ contribute essentially to the integral. Under the assumption that F_{mk} and F_{mn} vary only slowly with m at each location, one will then get:

$$\sum_m |a_m|^2 = \sum_m \left| F_{mk} - \sum_n \frac{F_{mn}\,F_{nk}}{E_n - E_k} \right|^2 \cdot \frac{4\pi^2\,t}{h \cdot \Delta E_m}\,, \tag{44}$$

which is the total probability of occupying the state whose energy equals E_m (is approximately equal to the energy E_k of the initial state, resp.) at time t. Contributing to that are, first of all, a direct transition (F_{mk}) and secondly, an indirect transition ($F_{mn} \cdot F_{nk}$) by way of a detour through intermediate states.

We now imagine that calculation as being performed on equation (32) and get:

$$\sum_{N_n} |\Phi_N^n(t)|^2 = \left| \mathfrak{F}_{NM}^{nm} - \sum_{Q_q} \frac{\mathfrak{F}_{NM}^{nq}\;\mathfrak{F}_{NM}^{qm}}{E_Q^q - E_M^m} \right|^2 \cdot \frac{4\pi^2\,t}{h \cdot \Delta E_N^n} \tag{45}$$

as the total probability of occupying states of the system of light + electron whose energies are equal to $E_N^n = E_N + E_n$ (almost equal to the energy $E_M^m = E_M + E_m$ of the initial state, resp.) at time t. The values of $\mathfrak{F}_{NM}^{nm} = \mathfrak{H}_{NM}^{nm}$ can now be once more be read off from (29).

We take the initial state to be M and the final state to be N in (39), but with $N_l = 0$ (scattering of ν_k-light into the previously radiation-free state ν_k) so:

$$\mathfrak{F}_{NM}^{nm} = \mathfrak{H}_{NM}^{nm} = \frac{\varepsilon^2 h \cos\alpha_{kl}}{2\pi\,\mu V\sqrt{\nu_k\,\nu_l}}\sqrt{N_k} \qquad \text{and} \qquad \Delta E_N^n = h\,\Delta\nu_l\,. \tag{46}$$

As an intermediate state $\begin{pmatrix} q \\ Q \end{pmatrix}$, we shall employ a state of the atom q such that the transition $\begin{pmatrix} q\,m \\ Q\,M \end{pmatrix}$ is the *absorption* of a ν_k and the transition $\begin{pmatrix} n\,q \\ N\,Q \end{pmatrix}$ is the *emission* of a ν_l. According to (34′) and (37), that will belong to:

$$\left.\begin{aligned} \mathfrak{H}_{QM}^{qm} &= \sqrt{\frac{h}{2\pi V\nu_k}}\,\dot{\mathfrak{M}}_k^{qm}\sqrt{N_k}\,, & \mathfrak{H}_{NQ}^{nq} &= \sqrt{\frac{h}{2\pi V\nu_l}}\,\dot{\mathfrak{M}}_l^{nq}\,, \\ E_Q^q - E_M^m &= h(\nu_{mq}-\nu_k)\,, & E_N^n - E_Q^q &= h(\nu_{qn}+\nu_l)\,, \end{aligned}\right\} \qquad (46')$$

in which q must be such that:

$$E_N^n - E_M^m = h\,(\nu_{mq} + \nu_{qn} + \nu_l - \nu_k) \approx 0\,.$$

We shall not consider further intermediate states. With (46), (46′), (46″), (45) will imply that:

$$\sum_{N_n} |\,\Phi_N^n(t)\,|^2 = \frac{4\pi^2 t}{h\cdot\Delta\nu_l}\,N_k\cdot\frac{1}{4\pi^2 V^2\nu_l\,\nu_k}\left|\frac{h\,\varepsilon^2}{\mu}\cos\alpha_{kl} - \sum\left(\frac{\dot{\mathfrak{M}}_l^{qm}\,\mathfrak{M}_k^{nq}}{\nu_{mq}+\nu_l} + \frac{\dot{\mathfrak{M}}_k^{qm}\,\mathfrak{M}_l^{nq}}{\nu_{mq}-\nu_l}\right)\right|^2 \qquad (47)$$

is the total probability of occupying state whose energies E_N^n lie close to the initial energy E_M^m at time t, i.e., the *total probability for the conversion of a light quantum* ν_k *into* ν_l at the expense of an atomic jump.

One will come to *dispersion* in the process for which the vanishing and newly-created light quantum indeed possess different direction, but almost the same frequency ($\nu_k = \nu_l$). Due to the fact that $E_N^n - E_M^m = 0$, that will then imply that the atomic states n and m are identical ($E^n = E^m$). In that way, when one writes $\nu_k = \nu_l = n$ and multiplies by $\sigma_k\,\sigma_l = V^2\cdot(\nu^4/c^6)\,\Delta\nu_l\cdot\Delta\omega_l\cdot\Delta\nu_k\cdot\Delta\omega_k$ [cf., (25′)], (47) will imply that:

$$N_k\,t\,\frac{\nu^2}{h^2c^5}\,\Delta\omega_l\,\Delta\nu_k\,\Delta\omega_k\cdot\left|\frac{h\,\varepsilon^2}{\mu}\cos\alpha_{kl} - \sum\left(\frac{\dot{\mathfrak{M}}_l^{qm}\,\mathfrak{M}_k^{nq}}{\nu_{nq}+\nu} + \frac{\dot{\mathfrak{M}}_k^{qm}\,\mathfrak{M}_l^{nq}}{\nu_{nq}-\nu}\right)\right|^2 \qquad (48)$$

is equal to the number of light quanta that leave the ray bundle $\Delta\nu\,\Delta\omega_k$ and are scattered into the cone $\Delta\omega_l$ during t.

One can formally reduce the creation of scattered light quanta to the emission of scattered light by a classical dipole of moment $\mathfrak{M}$ with the component $\mathfrak{M}_l$ in the direction of the electric vector. The energy that is emitted from the dipole into the cone of angles $\Delta\omega_l$ during time t is now:

$$\Delta\omega_l\cdot\frac{8\pi^3\nu^4\,\mathfrak{M}_l^2\,t}{c^3} \qquad \text{[cf., (36)],}$$

so the number of light quanta $h\nu$ that are emitted by it will be:

$$\Delta\omega_l \cdot \frac{8\pi^3 \nu^3 \mathfrak{M}_l^2 t}{h c^3} .$$

When that is set equal to (48), that will give the same effect as creating a classical dipole moment $\mathfrak{M}_l = \varepsilon \, \mathfrak{r}_l$:

$$\mathfrak{M}_l = \sqrt{\frac{N_k \, \Delta\nu_k \, \Delta\omega_k}{8\pi^3 h \nu c^3}} \; \varepsilon^2 \cdot \left| \frac{h}{\mu} \cos\alpha_{kl} - \sum \left(\frac{\dot{\mathfrak{r}}_l^{qm} \, \dot{\mathfrak{r}}_k^{nq}}{\nu_{nq} + \nu} + \frac{\dot{\mathfrak{r}}_k^{qm} \, \dot{\mathfrak{r}}_l^{nq}}{\nu_{nq} - \nu} \right) \right|^2 . \tag{49}$$

One can simplify that expression. First of all [cf., (40′)], the field strength amplitude that belongs to the ray bundle $N_k \, \Delta\nu_k \, \Delta\omega_k$ is:

$$\frac{c}{8\pi} \mathfrak{E}_k^2 = \mathfrak{K}(\nu) \, \Delta\omega_k \, \Delta\nu_k = \frac{N_k \, h \nu^3}{c^2} \Delta\omega_k \, \Delta\nu_k ,$$

so

$$\mathfrak{M}_l = \frac{\mathfrak{E}_k \varepsilon^2}{h} \frac{1}{8\pi^2 \nu^2} \cdot \left| \cdots \right| .$$

According to **Dirac**, the absolute value | … | can also be converted when one sets $\dot{\mathfrak{r}}^{qn} = 2i\pi\nu_{qn} \, \mathfrak{r}^{qn}$ and relates cos α_{kl} to $\mathfrak{r}_k^{qn}$, $\mathfrak{r}_l^{qn}$, and ν_{qn} and ν. **Dirac** ultimately obtained:

$$\mathfrak{M}_l = \frac{\mathfrak{E}_k \varepsilon^2}{h} \cdot \left| \sum_q \left(\frac{\mathfrak{r}_l^{qn} \, \mathfrak{r}_k^{nq}}{\nu_{nq} + \nu} + \frac{\mathfrak{r}_k^{qn} \, \mathfrak{r}_l^{nq}}{\nu_{nq} - \nu} \right) \right| ,$$

and when one replaces the vector component $\mathfrak{M}_l$ with $\mathfrak{M}$ and the component $\mathfrak{r}_l$ with the vector $\mathfrak{r}$:

$$\mathfrak{M} = \frac{\mathfrak{E}_k \varepsilon^2}{h} \cdot \left| \sum_q \left(\frac{\mathfrak{r}^{qn} \, \mathfrak{r}_k^{nq}}{\nu_{nq} + \nu} + \frac{\mathfrak{r}_k^{qn} \, \mathfrak{r}^{nq}}{\nu_{nq} - \nu} \right) \right| \tag{50}$$

will be the dipole moment that gave rise to the scattering that was described in (48) classically. Now, (50) coincides with a dispersion formula that was present by **Kramers** and **Heisenberg** (1), in conjunction with **Ladenburg** on the basis of conclusions that followed from the correspondence principle. The closer that ν lies to one of the transition frequencies ν_{nq} of the atom from its state $n = m$ to an intermediate state q, the greater that $\mathfrak{M}$ will be. Indeed, maxima of $\mathfrak{M}$ will occur at

---

(1) **H. A. Kramers**, Nature **113** (1924), pp. 673. – **Kramers** and **Heisenberg**, Zeit. Phys. **31** (1925), pp. 681. Classically, $\mathfrak{r}^{nq}$ corresponds to the amplitude of the upper frequency $(n - q)$ ν_0 of an (anharmonic) oscillator.

locations $\nu = \nu_{nq}$ with positive ν_{nq} , as well as locations $\nu = -\ \nu_{nq}$ with negative ν_{nq} , which correspond to transitions in atoms from the state n to lower and higher states, respectively.

CHAPTER VI

RELATIVISTIC WAVE MECHANICS

51. Relativistic electron in a field (¹). – The relativistic electron requires special treatment, since its kinetic energy is not, as was assumed up to now, a quadratic function of the coordinates. In order to next look for the **Hamilton-Jacobi** equation for the electron in a field, we shall use the coordinates:

$$x_1, x_2, x_3, x_4 = x, y, z, i\,c\,t,$$

as is customary in the theory of relativity, and introduce the mass μ and the proper time increment $d\tau$ by:

$$\mu = \frac{\mu_0}{\sqrt{1-\beta^2}}, \qquad d\tau = dt\cdot\sqrt{1-\beta^2}, \qquad \beta = \frac{v}{c}, \qquad \mu_0 = \text{rest mass.} \tag{1}$$

If differentiation with respect to t is suggested by a dot then one will have the identity:

$$\frac{1}{2\mu_0}[(\mu\,\dot{x}_1)^2 + (\mu\,\dot{x}_2)^2 + (\mu\,\dot{x}_3)^2 + (\mu\,\dot{x}_4)^2] + \frac{1}{2}\mu_0\,c^2 \equiv 0\,. \tag{1'}$$

In the absence of an external field, the quantities $\mu\,\dot{x}_k$ will be *impulse* p_k, and the latter equation will take the form:

$$p_1^2 + p_2^2 + p_3^2 + p_4^2 + \mu_0^2\,c^2 = 0 \qquad \left(p_4 = -\frac{E}{i\,c},\ x_4 = i\,c\,t\right). \tag{1''}$$

Its wave-mechanical conversion $\left(p_k = \dfrac{h}{2i\pi}\dfrac{\partial}{\partial x_k}\right)$ will then be:

$$\sum_k \frac{\partial^2\psi}{\partial x_k^2} + \left(\frac{2i\pi}{h}\right)^2 \mu_0^2\,c^2\,\psi = 0\,, \qquad \Box\psi + \left(\frac{2i\pi}{h}\right)^2 \mu_0^2\,c^2\,\psi = 0\,,$$

and it possesses the particular solution (homogeneous plane wave):

$$\psi(x_1, x_2, x_3, x_4) = A e^{2i\pi(p_1 x_1 + \cdots + p_4 x_4)/h},$$

(¹) **E. Schrödinger**, "Quantisierung als Eigenwertproblem, 4. Mitteilung," Ann. Phys. (Leipzig) **81** (1926), pp. 109. – **O. Klein** "Elektrodynamik und Wellenmechanik vom Standpunkt des Korrespondenzprinzips," Zeit. Phys. **41** (1927), pp. 407. – **W. Gordon**, "Der Comptoneffekt nach der **Schrödingerschen** Theorie," Zeit. Phys. **40** (1926), pp. 177. – **P. A. M. Dirac**, Proc. Roy. Soc. **111** (1926), pp. 405.

in which the otherwise-arbitrary constants p_k have to satisfy the condition (1″). One gets the general solution ψ by summing over the various particular solutions that belong to the different systems of values for the constants (p_1, p_2, p_3, p_4), but in each case, the p_k must satisfy the condition (1″):

$$\psi = \sum A(p_1, p_2, p_3, p_4)\, e^{2i\pi(p_1 x_1 + \cdots + p_4 x_4)/h} .$$

The coefficients A (p_1, p_2, p_3, p_4) are the *probability amplitudes* for the existence of an electron with the impulse values p_1, p_2, p_3, p_4 .

In the presence of a field, one once more first seeks the point-mechanical generalization of (1′) and then reinterprets that wave-mechanically. The electromagnetic field is derived from a four-vector Φ (potential) with the components:

$$\Phi_1, \Phi_2, \Phi_3, \Phi_4 = \mathfrak{A}_x, \mathfrak{A}_y, \mathfrak{A}_z, i\,\varphi \qquad \text{(vector and scalar potential).}$$

The *six-vector F* (viz., the field) is defined by:

$$F_{ik} = \frac{\partial \Phi_k}{\partial x_i} - \frac{\partial \Phi_i}{\partial x_k} = -F_{ki} , \qquad \text{or briefly} \qquad \text{Rot}\,\Phi = F . \tag{2}$$

The customary notation for the field components is:

$$F_{23}, F_{31}, F_{12}, F_{41}, F_{42}, F_{43} = \mathfrak{H}_x, \mathfrak{H}_y, \mathfrak{H}_z, i\,\mathfrak{E}_x, i\,\mathfrak{E}_y, i\,\mathfrak{E}_z,$$

and (2) means that:

$$\mathfrak{H} = \text{rot}\,\mathfrak{A}, \qquad \mathfrak{E} = -\,\text{grad}\,\varphi - \frac{1}{c}\dot{\mathfrak{A}} .$$

As an auxiliary condition for Φ, one takes:

$$\frac{\partial \Phi_1}{\partial x_1} + \cdots + \frac{\partial \Phi_4}{\partial x_4} = 0 , \qquad \text{or briefly} \qquad \text{Div}\,\Phi = 0 . \tag{2′}$$

The equations of motion of the electron:

$$\frac{d}{d\tau}\left(\mu_0 \frac{dx_k}{d\tau}\right) = -\,\frac{\varepsilon}{c}\sum_i \frac{dx_k}{d\tau}\,\text{Rot}_{ik}\Phi , \qquad \text{or briefly} \qquad \frac{d}{d\tau}\left(\mu_0 \frac{dx_k}{d\tau}\right) = -\,\frac{\varepsilon}{c}\left[\frac{dx_k}{d\tau}, \text{Rot}\,\Phi\right]$$

will take the canonical form:

$$\frac{dx_k}{d\tau} = \frac{\partial H}{\partial p_k}, \qquad \frac{dp_k}{d\tau} = -\,\frac{\partial H}{\partial x_k}$$

when one understands H (x, p) to mean the function:

$$H = \frac{1}{2\mu_0}\left[\left(p_1+\frac{\varepsilon}{c}\Phi_1\right)^2+\cdots+\left(p_4+\frac{\varepsilon}{c}\Phi_4\right)^2\right]+\frac{1}{2}\mu_0\,c^2, \tag{3}$$

which emerges from the left-hand side of the identity (1′) by the replacement:

$$\mu_0\frac{dx_k}{d\tau} = \mu\,\dot{x}_k = p_k + \frac{\varepsilon}{c}\Phi_k\,. \tag{4}$$

The identity (1′) then assume the form $H\ (x,\ p) = 0$. One then obtains the **Hamilton-Jacobi** equation from that by introducing an action function $S\ (x_1,\ \ldots,\ x_4)$ with $p_k = \partial S\ /\ \partial x_k$ in the form:

$$H\left(x,\frac{\partial S}{\partial x}\right)=\frac{1}{2\mu_0}\left\{\left(\frac{\partial S}{\partial x_1}+\frac{\varepsilon}{c}\Phi_1\right)^2+\cdots\right\}+\frac{1}{2}\mu_0\,c^2=0$$

or more briefly

$$\frac{1}{2\mu_0}\left(\text{grad}\,S+\frac{\varepsilon}{c}\Phi\right)^2+\frac{1}{2}\mu_0\,c^2=0. \tag{4′}$$

The **de Broglie** theory constructs the wave function $\psi = e^{2i\pi S/h}$ from a complete solution S of *that* equation.

The transition to **Schrödinger**'s wave mechanics will come about then one replaces p_k, not with $\partial S\ /\ \partial x_k$, but with the operator $\dfrac{h}{2i\pi}\dfrac{\partial}{\partial x_k}$, and constructs the equation:

$$\boxed{\left\{\frac{1}{2\mu_0}\left(\frac{h}{2i\,\pi}\text{Grad}+\frac{\varepsilon}{c}\Phi\right)^2+\frac{1}{2}\mu_0\,c^2,\psi\right\}=0,} \tag{5}$$

in place of $H\ (x,\ p)$. That equation for $\psi\ (x_1,\ x_2,\ x_3,\ x_4)$, which several authors (1) posed simultaneously, can also be obtained, as **Gordon** showed (2), as the **Euler** equation ($\tilde{\psi}$ = conjugate of ψ, $\psi_{x_k} = \dfrac{\partial\psi}{\partial x_k}$, $\tilde{\psi}_{x_k} = \dfrac{\partial\tilde{\psi}}{\partial x_k}$):

$$\sum_k\frac{d}{dx_k}\left(\frac{\partial L}{\partial\tilde{\psi}_{x_k}}\right)-\frac{\partial L}{\partial\tilde{\psi}} = 0$$

for the **Lagrangian** function:

---

(1) **E. Schrödinger**, Ann. Phys. (Leipzig) **81** (1926), pp. 109. – **V. Fock**, Zeit. Phys. **39** (1926), pp. 226. – **J. Kudar**, Ann. Phys. (Leipzig) **81** (1926), pp. 632.

(2) **W. Gordon**, further see **H. Bateman**, Proc. Nat. Acad. Amer. **13** (1927), pp. 326; Phys. Rev. **30** (1927), pp. 55. – **E. Guth**, Zeit. Phys. **41** (1927), pp. 235.

$$L = \sum_k \left(\psi_{x_k} + \frac{2i\pi}{h}\frac{\varepsilon}{c}\Phi_k\psi\right)\left(\tilde{\psi}_{x_k} - \frac{2i\pi}{h}\frac{\varepsilon}{c}\Phi_k\tilde{\psi}\right) + \left(\frac{2\pi\mu_0 c}{h}\right)^2 \psi\tilde{\psi} \ .$$

When one goes to the conjugates, x_4 is always treated like a real quantity in so doing, so one does not, say, set $\tilde{x}_4 = -x_4$.

In detail, when one observes (2′), (5) will read:

$$\Box\psi + \frac{2i\pi\varepsilon}{hc} 2\Phi \,\mathrm{Grad}\,\psi + \left[\left(\frac{2i\pi}{h}\frac{\varepsilon}{c}\right)^2 \Phi^2 + \left(\frac{2i\pi}{h}\right)^2 \mu_0^2 c^2\right]\psi = 0 \ . \tag{6}$$

If one introduces the special solution Ansatz $\psi = e^{2i\pi S/h}$ then the defining equation for S will be true:

$$\frac{h}{2i\pi}\Box S + \left(\mathrm{Grad}\, S + \frac{\varepsilon}{c}\Phi\right)^2 + \mu_0^2 c^2 = 0 \ ,$$

which differs from the classical equation (4) by only the first term, which is proportional to h.

By contrast, in optics of the light quantum equation:

$$p_1^2 + p_2^2 + p_3^2 + p_4^2 = 0 \qquad \left(p_4 = \frac{iE}{c} = \frac{ih\nu}{c}, \sqrt{p_1^2 + p_2^2 + p_3^2} = \frac{h\nu}{c}\right) ,$$

when one replaces p_k with $\frac{h}{2i\pi}\frac{\partial}{\partial x_k}$, the wave equation will become $\Box\psi = 0$, which does not include h.

We shall consider some special cases.

a) *In the absence of a force field* ($\Phi = 0$), one will get from (6):

$$\Box\psi + \left(\frac{2i\pi}{h}\right)^2 \mu_0^2 c^2 \psi = 0 \ . \tag{6′}$$

b) *In a static force field*, one introduces the Ansatz $\psi = \psi_n e^{2i\pi\nu_n t}$ in (6), in which ψ_n no longer depends upon t, and ν_n means a constant. One will then obtain the equation for ψ_n :

$$\Delta\psi_n + \frac{2i\pi\varepsilon}{hc} 2\mathfrak{A}\,\mathrm{grad}\,\psi + \left(\frac{2i\pi}{h}\right)^2 \left[\frac{\varepsilon^2}{c^2}\mathfrak{A}^2 - \frac{1}{c^2}(h\nu_n - \varepsilon\varphi)^2 + \mu_0^2 c^2\right]\cdot\psi_n = 0 \ , \tag{6″}$$

with the eigenvalues $h\nu_n$ for the energy and the normalized eigen-solutions ψ_n .

c) When the solution Ansatz $\psi = \psi^*(x_1, x_2, x_3, x_4)\cdot e^{-2i\pi\mu_0 c^2 t/h}$ is introduced into (6), that will give the following equation for ψ^*:

$$\Delta\psi^* - \frac{1}{c^2}\frac{\partial^2\psi^*}{\partial t^2} - \frac{8\pi^2\mu_0}{h^2}\,\varepsilon\varphi\cdot\psi^* + \frac{4i\pi}{h}\left\{\frac{\varepsilon}{c}\mathfrak{A}\,\text{grad} + \frac{\varepsilon\varphi}{c^2}\frac{\partial}{\partial t} - \mu_0, \psi^*\right\} + \left(\frac{2i\pi}{h}\cdot\frac{\varepsilon}{c}\right)^2(\mathfrak{A}^2 - \varphi^2)\psi^* = 0 .$$

If one neglects the terms that possess the factor $1/c^2$ or include the field potential quadratically then the will give the *non-relativistic* equation:

$$\Delta\psi^* - \frac{8\pi^2\mu_0}{h^2}\,\varepsilon\varphi\cdot\psi^* + \frac{4i\pi}{h}\left(\frac{\varepsilon}{c}\mathfrak{A}\,\text{grad}\,\psi^* - \mu_0\frac{\partial\psi^*}{\partial t}\right) = 0 , \qquad (6'')$$

which represents the generalization of the non-relativistic **Schrödinger** equation in the presence of a *magnetic field.*

52. Hydrodynamical interpretation and radiation. ([1]). – One can associate every eigensolution ψ of the wave equation (6) with a space-time flow field that is given by a current vector $\mathfrak{J}$ with the components $\mathfrak{J}_1, \ldots, \mathfrak{J}_4$ that fulfills the continuity equation Div $\mathfrak{J} = 0$. One multiplies (6) by $\tilde{\psi}$ and the equation conjugate to (6) by ψ and subtracts them. Due to the fact that Div $\Phi = 0$, the equation that arises:

$$(\tilde{\psi}\,\Box\psi - \psi\,\Box\tilde{\psi}) + \frac{2i\pi\varepsilon}{hc}\,2\,\Phi\,(\tilde{\psi}\,\text{Grad}\,\psi + \psi\,\text{Grad}\,\tilde{\psi}) = 0$$

can then be put into the form:

$$\text{Div}\,\mathfrak{J} = 0 \qquad \text{with} \qquad \mathfrak{J} = \frac{h}{2i\pi}\left(\tilde{\psi}\,\text{Grad}\,\psi + \psi\,\text{Grad}\,\tilde{\psi} + \frac{4i\pi\varepsilon}{hc}\Phi\,\psi\,\tilde{\psi}\right) . \qquad (7)$$

One gets the usual form div $(\rho\,\mathfrak{v}) + \partial\rho/\partial t = 0$ for the continuity equation from Div $\mathfrak{J} = 0$ from the relation:

$$\mathfrak{J}_1 = \rho\,\mathfrak{v}_x , \qquad \mathfrak{J}_2 = \rho\,\mathfrak{v}_y , \qquad \mathfrak{J}_3 = \rho\,\mathfrak{v}_z , \qquad \mathfrak{J}_4 = i\,c\,\rho .$$

---

[1] **W. Gordon**, Zeit. Phys. **40** (1926), pp. 117.

Upon introducing the rest density $\rho_0 = \rho \cdot \sqrt{1-\frac{v^2}{c^2}} = \frac{i}{c}|\mathfrak{J}|$ and the proper-time differential $d\tau = dt \cdot \sqrt{1-\frac{v^2}{c^2}}$, one will get the *four-velocity* $\mathfrak{V}$ with the components $\mathfrak{V}_k = dx_k / d\tau$:

$$\mathfrak{V}_1 = \frac{dx_1}{d\tau} = \frac{\mathfrak{v}_x}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{\mathfrak{J}_1}{\rho_0}, \quad \ldots, \qquad \mathfrak{V}_4 = i\,c\frac{dt}{d\tau} = \frac{i\,c}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{\mathfrak{J}_4}{\rho_0}, \tag{8}$$

$$\sum \mathfrak{V}_k^2 = -\,c^2, \qquad\qquad \mathfrak{J}^2 = -\,c^2\rho_0^2. \tag{9}$$

In a static field, one composes the general solution of (6) from the eigen-solutions of (6′):

$$\psi(x_1, \ldots, x_4) = \sum_n b_n\,\psi_n(x_1, x_2, x_3) \cdot e^{2i\pi\nu_n t} \tag{10}$$

with arbitrary constant coefficients b_n. The current vector $\mathfrak{J}$ can then be written in the form:

$$\left.\begin{aligned} &\mathfrak{J} = \sum_l \sum_m b_l\, b_m\, \mathfrak{J}^{(lm)} \\ \text{in which:}\qquad &\mathfrak{J}^{(lm)} = \frac{h}{4i\,\pi}\left(\tilde{\psi}_m \operatorname{Grad}\psi_l - \psi_l \operatorname{Grad}\tilde{\psi}_m + \frac{4i\pi\,\varepsilon}{h\,c}\Phi\,\psi_l\,\tilde{\psi}_m\right), \end{aligned}\right\} \tag{11}$$

corresponding to (7), which has the components:

$$\left.\begin{aligned} \mathfrak{J}_1^{(ml)} &= (\rho\,\mathfrak{v}_x)^{lm} = \frac{h}{4i\,\pi}\left(\tilde{\psi}_m \frac{\partial\psi_l}{\partial x_1} - \psi_l \frac{\partial\tilde{\psi}_m}{\partial x_1} + \frac{4i\,\pi\,\varepsilon}{h\,c}\Phi\,\psi_l\,\tilde{\psi}_m\right) \cdot e^{2i\pi(\nu_l-\nu_m)t}, \\ \mathfrak{J}_4^{(ml)} &= i\,c\,\rho^{lm} = \left(\frac{h\nu_l + h\nu_m}{2i\,c} + \frac{\varepsilon}{c}\Phi\right)\psi_l\,\tilde{\psi}_m \cdot e^{2i\pi(\nu_l-\nu_m)t}. \end{aligned}\right\} \tag{12}$$

Just as one did with the "transition density" $\rho^{(lm)}$ in no. **30**, so can one introduce the "transition current" $\mathfrak{J}^{(lm)}$ here as the source of electromagnetic radiation (**Gordon**): When the electron jumps from the state l to the state m in the field Φ, the current field $\mathfrak{J}^{(lm)}$ shall be regarded as the source of a secondary radiation field with the retarded potential Φ^*:

$$\Phi^* = \frac{1}{c}\int \frac{[\mathfrak{J}^{(lm)}]_{t-r/c}}{r}\,dx_1\,dx_2\,dx_3\,, \tag{12′}$$

as in **Maxwell**’s theory. On the basis of that assumption, **W. Gordon** succeeded in calculating the frequency and intensity of the secondary light field Φ^* that is radiated during the transition of a free electron from the state of motion l to m in the field of a primary light source Φ (**Compton** scattering of light). Moreover, one sees from (12), (12′) that under the transition lm, the secondary field Φ^* will possess the temporal period $\nu_{lm} = \nu_l - \nu_m = (E_l - E_m) / h$, which corresponds to the *combination principle* and the **Bohr** frequency condition.

It is interesting to introduce the Ansatz that was used before (no. **37**):

$$\psi = \alpha(x_1,\dots,x_4)\cdot e^{2i\pi S(x_1,\dots,x_4)/h}, \qquad \text{so} \qquad \alpha^2 = \psi\tilde{\psi}, \qquad S = \frac{h}{4i\pi}\ln\left(\frac{\psi}{\tilde{\psi}}\right) \tag{13}$$

into the wave equation (6) (α and S are real). Upon separating the real and imaginary parts, that will give *two* equations for α and S:

$$\left.\begin{aligned} &\frac{h^2}{4\pi^2}\Box\alpha - \alpha(\mathrm{Grad}\,S)^2 - 2\alpha\frac{\varepsilon}{c}\Phi\,\mathrm{Grad}\,S - \alpha\cdot\left(\frac{\varepsilon^2\,\Phi^2}{c^2} + \mu_0^2\,c^2\right) = 0,\\ &\alpha\Box S + 2\,\mathrm{Grad}\,\alpha\cdot\mathrm{Grad}\,S + 2\frac{\varepsilon\Phi}{c}\mathrm{Grad}\,\alpha = 0.\end{aligned}\right\} \tag{13′}$$

Due to the fact that Div $\Phi = 0$, the second one can also be written in the form:

$$\mathrm{Div}\,\mathfrak{J} = 0, \qquad \text{with} \qquad \mathfrak{J} = \alpha^2\left(\mathrm{Grad}\,S + \frac{\varepsilon}{c}\Phi\right), \tag{14}$$

while the first one can be written in the form:

$$\frac{1}{2\mu_0}\left(\mathrm{Grad}\,S + \frac{\varepsilon}{c}\Phi\right)^2 + \frac{1}{2}\mu_0\,c^2 - \frac{h^2}{8\pi^2\mu_0}\frac{\Box\alpha}{\alpha} = 0\,. \tag{14′}$$

That should be compared to the classical equation (4′). Due to (8), one has:

$$\rho_0^2 = -\frac{\mathfrak{J}^2}{c^2}\,, \tag{14″}$$

and from (14), (14′), the four-velocity $\mathfrak{B} = \mathfrak{J} / \rho_0$ will then become:

$$\mathfrak{B} = \frac{\alpha^2}{\rho_0}\left(\text{Grad}\, S + \frac{\varepsilon}{c}\Phi\right) \qquad \text{with} \qquad \rho_0 = \sqrt{-\frac{\mathfrak{J}^2}{c^2}} = \mu_0\,\alpha^2\sqrt{1-\frac{h^2}{4\pi^2\mu_0^2\,c^2}\cdot\frac{\Box\alpha}{\alpha}}\,. \tag{15}$$

We now seek the *equations of motion* of the current, by analogy with **Madelung**'s hydrodynamical equations of motion (no. **37**) for non-relativistic undulatory mechanics. In order to do that, we start from the identity for an arbitrary four-vector:

$$\frac{d\mathfrak{B}}{d\tau} = (\mathfrak{B}\ \text{Grad})\ \mathfrak{B} \quad \text{and} \quad \tfrac{1}{2}\text{Grad}\ \mathfrak{B}^2 = (\mathfrak{B}\ \text{Grad})\ \mathfrak{B} + [\mathfrak{B},\ \text{Rot}\ \mathfrak{B}],$$

in which $d / d\tau$ described the change in $\mathfrak{B}$ when one advances along the flow. In particular, for $\mathfrak{B} = \mathfrak{B}\ \mu_0$, one will have:

$$\frac{d}{d\tau}(\mu_0\,\mathfrak{B}) = \mu_0\,(\mathfrak{B}\ \text{Grad})\ \mathfrak{B} = -\,\mu_0\,[\mathfrak{B},\ \text{Rot}\ \mathfrak{B}]\,,$$

since $\tfrac{1}{2}\text{Grad}\ \mathfrak{B}^2 = \tfrac{1}{2}\text{Grad}\ (-c^2) = 0$. One can then write the latter as:

$$\frac{d}{d\tau}(\mu_0\,\mathfrak{B}) = -\left[\mathfrak{B},\text{Rot}\,\frac{\mathfrak{J}}{\rho_0}\right] = -\left[\mathfrak{B},\text{Rot}\,\frac{\mathfrak{J}}{\alpha^2}\right] - \left[\mathfrak{B},\text{Rot}\left(\frac{\mathfrak{J}\,\mu_0}{\rho_0} - \frac{\mathfrak{J}}{\alpha^2}\right)\right],$$

and finally, since $\text{Rot}\left(\frac{\mathfrak{J}}{\alpha^2}\right) = \text{Rot}\left(\text{Grad}\, S + \frac{\varepsilon}{c}\Phi\right) = \frac{\varepsilon}{c}\ \text{Rot}\ \Phi$, one will have:

$$\frac{d}{d\tau}(\mathfrak{B}) = = -\left[\mathfrak{B},\text{Rot}\,\frac{\varepsilon}{\mu_0\,c}\Phi\right] - \left[\mathfrak{B},\text{Rot}\left(\frac{\mathfrak{J}}{\rho_0} - \frac{\mathfrak{J}}{\mu_0\,\alpha^2}\right)\right] = \mathfrak{K}_a + \mathfrak{K}_i = -\left[\mathfrak{B},\text{Rot}\,\frac{\varepsilon}{\mu_0\,c}\Phi\right] \tag{16}$$

for the force per unit mass. The potential that acts on the current is given by:

$$\Phi' = \Phi + \mathfrak{J}\left(\frac{\mu_0\,c}{\rho_0\,\varepsilon} - \frac{c}{\alpha^2\,\varepsilon}\right) = \left(\Phi + \frac{\mathfrak{B}}{c}\Phi_5\right),$$

when one understands Φ_5 to mean the scalar quantity:

$$\Phi_5 = \frac{c^2\mu_0}{\varepsilon}\left(1 - \frac{\rho_0}{\mu_0\,\alpha^2}\right). \tag{16′}$$

In the classical electrodynamics of charge fluid matter, and "external field" does not appear as the applied force per unit mass, but the force:

$$\mathfrak{K} = -\left[\mathfrak{B}, \mathrm{Rot}\,\frac{\varepsilon}{\mu_0\, c}\,\Phi\right]$$

per *unit mass*, in which the six-vector field Rot $\Phi = F$ is coupled with the current density by **Maxwell**'s equations with the current density $\mathfrak{J}$:

$$\Delta\mathrm{iv}\,F = \mathfrak{J}, \qquad \Delta\mathrm{iv}\;F^* = 0\,, \quad (F_{12}^* = F_{24}\text{ , etc.}),$$

and the charge per unit mass is given by ε / μ_0 . By contrast, in the context of the wave-mechanical current, nothing is said about that or any similar coupling of $\mathfrak{J}$ and Φ, but rather, Φ is the *external* field potential in that case, which is *independent* of the current, and it is extended to an applied potential Φ by an "internal stress potential" $(\mathfrak{B} / c)\,\Phi_5$.

One will get a *first approximation* for the quantum-theoretic current for $\mathfrak{v}^2 \ll c^2$ (i.e., when one neglects relativistic effects) when one uses the approximation:

$$\frac{\rho_0}{\mu_0\,\alpha^2} = 1 - \frac{h^2\cdot\Delta\alpha}{8\pi^2\,\mu_0^2\,c^2\cdot\alpha}\,,$$

instead of (15), while neglecting $\frac{1}{c^2}\ddot{\alpha}$ in comparison to $\Delta\alpha$ in $\Box\alpha$. In that approximation, one will get the equation of motion as:

$$\frac{d}{d\tau}\mathfrak{B} = -\left[\mathfrak{B}, \mathrm{Rot}\,\frac{\varepsilon}{\mu_0\, c}\,\Phi\right] + \mathfrak{K}_i \qquad \text{with} \qquad \mathfrak{K}_i = \frac{h^2}{8\pi^2\,\mu_0}\,\mathrm{grad}\left(\frac{\Delta\alpha}{\alpha}\right)$$

for the three spatial components of $\mathfrak{K}_i$, while the fourth component of $\mathfrak{K}_i$ will vanish completely in that approximation. The additional force $\mathfrak{K}_i$ in that then depends upon only the *density* α^2 and its spatial differential quotients and is identical to **Madelung**'s non-relativistic additional force [(23), in no. **37**] for the "internal" potential. Upon integrating the latter equation of motion over *space*, the integral of the additional force will vanish, precisely as it did in no. **38**, and what will result is a *center of mass theorem* that the center of mass of the streaming fluid ("wave packet") moves like a mechanical mass-point under the influence of only the external force that acts upon it.

By contrast, in a *higher approximation*, when one considers the relativistic terms, the internal force at each location will not only depend upon the *density distribution* of the current there, but also on its *velocity distribution*. As a result, the center of mass theorem will no longer be true, but rather, that additional force will also have an essential meaning for the motion of the wave packet as a whole. Therefore, whereas in the non-relativistic mechanics of the ψ-current, a wave packet will behave like a mass-point that moves according to classical mechanics, that relationship between mass-points and wave-packets will disappear in relativistic quantum mechanics.

53. Five-dimensional conception of wave mechanics (1). – One will get an especially symmetric form for wave mechanics of the electron by the process of **O. Klein**, **V. Fock** when one adds a fifth coordinate x_5 to the coordinates $x, y, z, i\,ct = x_1, x_2, x_3, x_4$ in a purely-formal way such that the wave function ψ will depend upon that fifth coordinate *periodically*. Namely, if one defines:

$$\psi(x_1, x_2, x_3, x_4, x_5) = \psi(x_1, x_2, x_3, x_4) \cdot e^{2i\pi c\mu_0 x_5}$$

then one can also write $\frac{\partial^n \psi}{\partial x_5^n}$ instead of $\left(\frac{2i\pi c\,\mu_0}{h}\right)^n \cdot \psi$ and in place of (6), one can now pose the equation for $\psi(x_1, \ldots, x_5)$:

$$\Box\psi + 2\frac{\varepsilon\Phi}{\mu_0 c^2}\frac{\partial}{\partial x_5}(\text{Grad}\,\psi) + \left[\left(\frac{\varepsilon\Phi}{\mu_0 c^2}\right)^2 + 1\right]\frac{\partial^2\psi}{\partial x_5^2} = 0\,. \tag{17}$$

In the absence of the field ($\Phi = 0$), (17) reduces to the form:

$$\boxed{\Delta_5\psi = 0,} \qquad \text{i.e.,} \qquad \left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{\partial^2\psi}{\partial x_5^2}\right) - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0 \tag{17′}$$

with the solution:

$$\psi = a\,e^{2i\pi(p_1 x_1 + \cdots + p_5 x_5)/h} \qquad \text{with} \qquad \boxed{\sum_{k=1}^{5} p_k^2 = 0,} \tag{17″}$$

or when one introduces the abbreviations:

$$p_1 = p_x\,, \qquad p_2 = p_y\,, \qquad p_3 = p_z\,, \qquad p_4 = \frac{iE}{c} = \frac{ih\nu}{c}\,, \qquad p_5 = \mu_0\,c\,,$$

$$\psi = a\,e^{2i\pi(-h\nu t + p_1 x_1 + \cdots + \mu_0 c x_5)/h}\,,$$

ψ will now represent plane wave in x, y, z, x_5-space whose phase propagates in time with a velocity v that is given by:

$$\frac{1}{v^2} = \left(\frac{p_x}{h\nu}\right)^2 + \left(\frac{p_y}{h\nu}\right)^2 + \left(\frac{p_z}{h\nu}\right)^2 + \left(\frac{\mu_0 c}{h\nu}\right)^2 = \frac{1}{c^2}\,, \qquad \text{i.e.} \qquad \boxed{v = c,} \tag{18}$$

i.e., with the universal velocity c, which is independent of $\nu = E / h$. The propagation in x, y, z, x_5-space is *dispersionless*. However, the trace of the plane-waves in x, y, z-space (x_5 = const.) propagates with the velocity u, which is given by:

---

(1) **O. Klein**, Zeit. Phys. **37** (1926), pp. 895; **V. Fock**, Zeit. Phys. **39** (1926), pp. 226; **P. Ehrenfest** and **G. Uhlenbeck**, Zeit. Phys. **39** (1926), pp. 495.

$$\frac{1}{u^2} = \left(\frac{p_x}{h\nu}\right)^2 + \left(\frac{p_y}{h\nu}\right)^2 + \left(\frac{p_z}{h\nu}\right)^2 = \frac{1}{c^2} - \frac{\mu_0^2\, c^2}{h^2\, \nu^2},$$

so:

$$u = \frac{c\, h\nu}{\sqrt{h^2\,\nu^2 - \mu_0^2\, c^4}} = \frac{h\nu}{\sqrt{-p_4^2 - p_5^2}} = \frac{h\nu}{\sqrt{p_x^2 + p_y^2 + p_z^2}} \,. \tag{18′}$$

u depends upon ν (**de Broglie** dispersion law), so the associated wavelength will be:

$$\lambda = \frac{u}{\nu} = \frac{h}{\sqrt{p_x^2 + p_y^2 + p_z^2}} = \frac{hc}{\sqrt{E^2 - \mu_0^2\, c^4}},$$

and it will always be real, since $\mu_0\, c^2$, as the rest energy, is always smaller than E.

If an *electromagnetic field* is present then we seek to solve the general equation (17) by the Ansatz:

$$\psi = \alpha\, e^{2i\pi S/h} = \alpha(x_1, \ldots, x_4)\, e^{2i\pi[S(x_1,\ldots,x_4) + c\,\mu_0\, x_5]/h} \qquad \text{with} \qquad \frac{\partial S}{\partial x_5} = c\,\mu_0 \,. \tag{19}$$

That will again lead to the two equations (13) for α and S [(14), (14′), resp.]. In particular, (14′) can then be written in the five-dimensional form:

$$\boxed{\left(\text{Grad}\, S + \frac{\varepsilon}{c}\Phi\right)^2 = 0,} \qquad \sum_{k=1}^{5}\left(\frac{\partial S}{\partial x_k} + \Phi_k\right)^2 = 0\,, \tag{19′}$$

when one, with **F. London** (¹), defines the scalar quantity (16′):

$$\frac{\varepsilon}{c}\Phi_5 = c\,\mu_0 - \frac{c\,\rho_0}{\alpha^2} = c\,\mu_0 - \sqrt{\mu_0^2\, c^2 - \frac{h^2}{4\pi^2}\frac{\Box\alpha}{\alpha}} \,.$$

Φ_5 then plays the role of a fifth electromagnetic potential component, along with Φ_1 to Φ_4, but in such a way that the components Φ_1 to Φ_4 depend upon the relativistic *reference system*, while Φ_5 is a scalar (i.e., invariant). Φ_5 is the relativistic counterpart to the **Madelung** potential of the internal forces of the ψ-field acting upon itself (no. **37**). If Φ_1 to $\Phi_4 = 0$ then α = const., $\Box\alpha = 0$, and $\Phi_5 = 0$.

One can eliminate the four components of the matter current $\mathfrak{J}_1, \ldots, \mathfrak{J}_4$ in (14) with the quantity:

(¹) **F. London**, Zeit. Phys. **42 (**1927), pp. 375.

$$\mathfrak{J}_5 = \alpha^2 \left(\frac{\partial S}{\partial x_5} + \frac{\varepsilon}{c} \Phi_5 \right) = \alpha^2 \sqrt{\mu_0^2 c^2 - \frac{h^2}{4\pi^2} \frac{\Box \alpha}{\alpha}} = c \rho_0 , \tag{19''}$$

such that (14″) can then be written in the form:

$$\sum_{k=1}^{5} \mathfrak{J}_k^2 = 0 , \qquad \mathfrak{J} = \alpha^2 \left(\text{Grad } S + \frac{\varepsilon}{c} \Phi \right). \tag{20}$$

The four-dimensional continuity equation also keeps the form Div $\mathfrak{J} = 0$ in five dimensions, since $\partial\mathfrak{J}_5 = / \partial x_5 = 0$, because a depends upon only $x_1, \ldots, x_4$. The flow velocity in the x_5-direction that belongs to $\mathfrak{J}_5$ will be described as an extension to the four components $\mathfrak{B}_1, \ldots, \mathfrak{B}_4$ of (15) by:

$$\mathfrak{B}_5 = \frac{dx_5}{d\tau} = c \left(\frac{\partial S}{\partial x_5} + \frac{\varepsilon}{c} \Phi_5 \right) : \sqrt{\mu_0^2 c^2 - \frac{h^2}{4\pi^2} \frac{\Box \alpha}{\alpha}} = c$$

with the help of (19′), and due to the fact that:

$$\mathfrak{B}_5 = c \qquad \text{and} \qquad \sum_{k=1}^{4} \mathfrak{B}_k^2 = - c^2 , \qquad \text{one will then have} \qquad \boxed{\sum_{k=1}^{5} \mathfrak{B}_k^2 = 0.} \tag{21}$$

The material points of the continuum will be such that, along with their temporal motion in x_1, x_2, x_3-space, they will also be assigned a certain advance in the x_5-direction:

$$dx_5 = c \cdot d\tau \qquad \left(\mathfrak{B}_5 = \frac{dx_5}{d\tau} = c \right). \tag{21'}$$

In that way, the streamlines in the $x_1 \ldots x_4$-world will belong to certain streamlines in the $x_1 \ldots x_5$-world.

A five-dimensional line-element $d\sigma$ that is parallel to the current $\mathfrak{J}$ will be "orthogonal to the current, because $\sum_{k=1}^{5} \mathfrak{J}_k^2 = 0$ means that $\mathfrak{J}$ is orthogonal to itself. One will then have:

$$\left. \begin{aligned} & (\mathfrak{J}\, d\sigma) = 0, \\ \text{and therefore, from (20):} \qquad & \\ & \text{Grad } S \cdot d\sigma = -\frac{\varepsilon}{c} \Phi \, d\sigma \end{aligned} \right\} \tag{22}$$

along a five-dimensional world-line element $d\sigma$.

If one follows $\psi = \alpha\, e^{2i\pi S/h} = e^{\ln\alpha + 2i\pi S/h}$ along a streamline then one will have:

$$d \ln \psi = \frac{d\psi}{\psi} = \left(\frac{d \ln\alpha}{d\sigma} + \frac{2i\pi}{h} \text{Grad } S \right) d\sigma ,$$

and with the use of (22):

$$\frac{d\psi}{\psi} = \sum_5 \left(\frac{\partial \ln\alpha}{\partial x_k} - \frac{2i\pi}{h} \frac{\varepsilon}{c} \Phi_k \right) dx_k \ . \tag{23}$$

Therefore, ψ will vary when one advances along the five-dimensional current according to the formula (¹):

$$\psi = \alpha\, e^{2i\pi S/h} = \psi_0 \exp\left[\frac{2i\pi}{h} \int \sum_5 \left(-\frac{\varepsilon}{c} \Phi_k + \frac{h}{2i\pi} \frac{\partial \ln\alpha}{\partial x_k} \right) dx_k \right] , \tag{24}$$

which is a result that be important in what follows. The formal advantage of the five-dimensional notation is expressed especially in the simplicity of the boxed formulas compared to the ones in the previous number. One should compare (17′) with (6′), (17″) with (1″), (18) with (18′), (19′) with (14′), and (21) with (9).

54. Weyl's theory of electromagnetism (²). – The relation concerning the change in the **Schrödinger** state function ψ along a streamline of the associated matter current that was just derived is significant in **Weyl**'s general-relativistic foundation of electromagnetism as a metric property of the space-time continuum (as an extension of **Einstein**'s explanation for gravitation). We might briefly summarize the foundations of **Einstein**'s and **Weyl**'s theories then. In the theory of general relativity, the space-time events are measured without referring to a special coordinate system in such a way that one associates neighboring events (two world-points with three spatial and one imaginary time coordinate) with the "world distance" *ds* according to a certain prescription: The measurement shall involve yardsticks and clocks that are appropriate to an inertial system (small free-falling boxes) and exist at the world-point in question. $(ds)^2$ will then be defined by the measurement result:

$$dX_1^2 + \cdots + dX_4^2 = dX^2 + dY^2 + dZ^2 - c^2 dT^2 = ds^2 , \tag{25}$$

in which *dX*, …, *dT* mean the readings in the inertial system. The yardsticks and clocks shall be calibrated in such a way that any two world-points of a light ray will possess the world distance $ds = 0$, so any yardstick will be taken to represent a unit of length and any clock to be a unit of time (which is how the numerical value of the constant *c* is first determined). The theory of relativity then states that the result of a measurement of *ds* for any two world-points in question will always

---

(¹) **F. London**, Zeit. Phys. **42** (1927), pp. 375.

(²) **H. Weyl**, Ber. Preuss. Akad. Wiss. (1918), pp. 465. Ann. Phys. (Leipzig) **59** (1919), pp. 101.

be the same, regardless of which of the many comoving inertial systems there are used to measure *ds*.

If the world-points are now described by *coordinates* $x_1, \ldots, x_4$ ($= i\,c\,t$) in an arbitrary non-inertial curvilinear coordinate system then the distance *ds* between two neighboring world-points that was measured above can also be expressed in terms of the associated coordinate differences in the form:

$$ds^2 = \sum_i \sum_k g^{ik}\, dx_i\, dx_k \tag{26}$$

with *metric coefficients* g^{ik} that depend upon the position (x_1, x_2, x_3, x_4) and can be determined by inertial measurements of *ds* in various directions at each world-location.

If one bases the description of the world-point on a different arbitrary coordinate system x_1', …, x_4' then the metric coefficients g'^{ik} in the measurement of *ds*:

$$ds^2 = \sum_i \sum_k g'^{ik}\, dx_i'\, dx_k'$$

in an inertial system and using the coordinate differences dx_k' between any two world-points would turn out differently.

Einstein's original theory initially assumed that the measurement was based, as it was in **Riemannian** geometry, in the fact that two distant line segments being measured could be compared with each other by transporting the unit yardsticks and unit clocks to distant world-points and regarding them as unchanged. However, **Weyl** considered that invariance of the calibration units under transport (**Riemannian** "geometry at a distance") to be a restrictive assumption that would be meaningful only if the forces (viz., electromagnetic fields) that prevail in the world-region where the measurements are being made are constant. By contrast, according to **Weyl**, one should generally expect a distortion of the yardsticks and clocks under transport due to the field in such a way that a new definition of the units of length and time would be required for any small world-region. If one would then wish to compare the segment *ds* with a distant segment ds^* at all then one must first make some assumptions about the change in the yardsticks and clocks under transport in the field. Now, **Weyl** arrived at a consistent "local geometry" by the assumption that the evolution of a material process could be described world-lengths *ds* of *different* magnitudes according to the position at which it took place, and indeed *ds* would change under a displacement ("parallel to itself") according to the linear formula:

$$\delta(ds^2) = ds^2 \cdot K \cdot \sum_k \varphi^k \delta x_k\,, \tag{26′}$$

in which K means a constant, and the coefficients φ^k are certain functions of the world-coordinates. One will then have:

$$\frac{\delta(ds^2)}{ds^2} = \delta \ln ds^2 = K \sum \varphi^k \delta x_k \ ,$$

and when one integrates that:

$$ds^2 = (ds^2)_0 \cdot e^{K \int \sum \varphi^k \, dx_k} , \tag{27}$$

one will have a formula for the change in ds^2 under finite parallel displacement. However, one can establish the differing quantity dS:

$$dS^2 = \frac{1}{\alpha} \cdot ds^2 \qquad \text{with} \qquad \delta(dS)^2 = dS^2 \cdot K \cdot \sum_k \Phi^k \delta x_k \tag{27′}$$

as the length of the line element for a *different* gauge, where α is an arbitrary coordinate function that characterizes the gauge, and the Φ^k are different functions accordingly.

On the contrary, the connection between the metric coefficients g^{ik} and G^{ik} for one and the same coordinate system, but different gauges, is implied by:

$$ds^2 = \sum_i \sum_k g^{ik} \, dx_i \cdot dx_k \qquad \text{and} \qquad dS^2 = \sum_i \sum_k G^{ik} \, dx_i \cdot dx_k$$

in the form:

$$G^{ik} = \frac{1}{\alpha} g^{ik} \qquad (\alpha = \text{gauge function}).$$

One will then get:

$$ds^2 = (ds^2)_0 \cdot \exp\left[K \int \sum \left(\Phi^k + \frac{1}{K} \frac{\partial \ln \alpha}{\partial x_k} \right) dx_k \right] \tag{29}$$

as the formula for the length ds of the line element under transport in that new gauge.

Now, in general, the change in ds^2 will depend upon the *path* of displacement, since ds^2 is not integrable. It is only when the integrand in the exponent is a complete differential, namely, when the rotation components:

$$F^{ik} = \text{Rot}^{ik}\, \Phi = \frac{\partial \Phi^k}{\partial x_i} - \frac{\partial \Phi^i}{\partial x_k} = - F^{ki} \tag{30}$$

vanish, that the change in ds^2 will be independent of the *path* along which the line element is taken from the starting point to the endpoint. Namely, $\frac{\partial}{\partial x_i} \frac{\partial \ln \alpha}{\partial x_k} - \frac{\partial}{\partial x_k} \frac{\partial \ln \alpha}{\partial x_i}$ (i.e., Rot Grad ln α) will vanish identically, so due to (28), one will always have:

$$\text{Rot}\, \Phi = \text{Rot}\, \varphi ,$$

independently of the arbitrary gauge function α. However, if $F^{ik} = 0$, but Φ^k is not equal to zero, then one can always choose the gauge function α in such a way that $\varphi^k = \Phi^k + \partial \ln \alpha / \partial x_k$ will vanish.

However, the Φ^k generally do not need to be arranged such that $F^{ik} = \mathrm{Rot}^{ik}\,\Phi$ vanishes. Nonetheless, one always has the identity:

$$\mathrm{Rot}^{ikl} F = \frac{\partial F^{ik}}{\partial x_l} + \frac{\partial F^{kl}}{\partial x_i} + \frac{\partial F^{li}}{\partial x_k} = 0 \qquad \text{for} \qquad F^{ik} = \mathrm{Rot}^{ik}\,\Phi . \tag{31}$$

However, those equations will go to precisely the ***Maxwell** equations* rot $\mathfrak{E} + \frac{1}{c}\dot{\mathfrak{H}} = 0$, div $\mathfrak{H} = 0$ of electrodynamics when one sets:

$$(F^{41}, F^{42}, F^{43}) = i\,\mathfrak{E}, \qquad (F^{23}, F^{31}, F^{12}) = \mathfrak{H}, \tag{31'}$$

i.e., Φ^k is identified with the electromagnetic potentials $\mathfrak{A}_x$, $\mathfrak{A}_y$, $\mathfrak{A}_z$, $i\,\varphi$, whose $\mathrm{Rot}^{ik}\,\Phi$ are known to determine the field strengths. The other pair of **Maxwell** equations rot $\mathfrak{H} - \frac{1}{c}\dot{\mathfrak{E}} = \frac{1}{c}\,j$, div $\mathfrak{E} = \rho$ can be written in the form:

$$\sum_k \frac{\partial F_{ik}}{\partial x_k} = \mathfrak{J}_i . \tag{32}$$

*In **Weyl**'s theory, the electromagnetic potentials are nothing but the metric coefficients Φ^k of the change in length of a line element under parallel displacement*. Moreover, due to the fact that $\mathrm{Rot}\,(\mathrm{Grad} \ln \alpha) \equiv 0$, the potential Φ^k will lead to the same field strengths as the potentials $\varphi^k = \Phi^k + \frac{1}{K}\frac{\partial \ln \alpha}{\partial x_k}$. Therefore, whereas the potentials are determined only up to the additive function Grad ln α, the *field strengths* are independent of the special choice of gauge function α. Just as in **Einstein**'s theory, the properties of gravitation (e.g., equality of inertial and gravitational mass) will become understandable when one attributes gravitation, not to forces in a Euclidian space, but to a deviation from the *Euclidian* metric, **Weyl**'s theory makes electromagnetic understandable when one attributes it, not to forces in a **Riemannian** space, but to a deviation from the **Riemannian** metric that comes from the variability of the magnitude of the units under transport (**Weyl**'s local geometry).

However, **Weyl**'s theory, in the general form above, next leads to a conflict with experiments (¹). For example, let $ds = \tau$ be the period of a clock that is at rest in an inertial system. In a constant pure electrostatic field ($\Phi^1 = \Phi^2 = \Phi^3 = 0$, $\Phi^4 = i\,\varphi = \text{const.}$), τ will change according to (27):

(¹) See, especially, **A. Einstein**, Ber. Preuss. Akad. Wiss. (1918), pp. 478.

$$\frac{\tau}{\tau_0} = e^{K/2\int i\varphi\cdot ic\,dt} = e^{-K\varphi\,c\,(t-t_0)/2} .$$

Now, if the clock is kept in the field only during the time $t - t_0$ then its period will *always* be greater or lesser (according to the sign of the constant K, which is assumed to be real) then the period of the same clock when it is always found in a zero field. In particular, that must manifest itself in an ongoing perturbation of the spectral lines of an atom under a temporary cessation of the field. Now, such an effect, which would prevent the appearance of sharp spectral lines, contradicts all experiments. However, the next number will refute that objection.

55. Periodicity of Weyl's mass along quantum paths. – **F. London** gave a way of getting around that contradiction in the form of a "remarkable property of quantum paths" (1) that **Schrödinger** found in the older **Bohr** theory: When the line integral extends over a spatially-closed quantum path:

$$\oint\sum_{k=1}^{4}\frac{\varepsilon}{c}\Phi^k dx_k = -\,n\,h\,, \tag{33}$$

it will give a whole-number multiple of **Planck**'s constant. With **F. London**, we can prove that equation, which **Schrödinger** showed in an example, relativistically as follows: According to (4), for the mechanical path of an electron ε, one will have:

$$\mu_0\frac{dx_k}{d\tau} = \frac{\partial S}{\partial x_k}+\frac{\varepsilon}{c}\Phi_k\,,$$

and therefore, from (4′):

$$\sum_{k=1}^{4}\left(\frac{\partial S}{\partial x_k}+\frac{\varepsilon}{c}\Phi_k\right)dx_k = -\,\mu_0\,c^2\,d\tau = -\,\mu_0\,c^2\sqrt{1-\frac{v^2}{c^2}}\,dt\ .$$

If one integrates that over a spatially-closed periodic quantum path then, as a result of the quantum condition:

$$\sum_{k=1}^{3}\oint\frac{\partial S}{\partial x_k}dx_k = n\,h\,,$$

one will get the relation:

$$\oint\frac{\partial S}{\partial x_4}dx_4+\oint\frac{\varepsilon}{c}\sum_{k=1}^{4}\Phi^k dx_k = -\,n\,h-\oint\mu_0\,c^2\sqrt{1-\frac{v^2}{c^2}}\,dt\,.$$

---

(1) **E. Schrödinger**, Zeit. Phys. **12** (1922), pp. 13.

Now, for a *quantum path* in a stationary electromagnetic field, the energy E of the electron is described by:

$$\frac{\partial S}{\partial t} = -E\,, \qquad \frac{\partial S}{\partial x_4}\,dx_4 = -E\,dt\,,$$

so as a result:

$$\oint \frac{\varepsilon}{c}\sum_{k=1}^{4}\Phi^k dx_k = -n\,h + \oint\left(-\mu_0\,c^2\sqrt{1-\frac{v^2}{c^2}}+E\right)dt\,. \tag{33′}$$

Now, the integral on the right-hand side vanishes. [That is because:

$$\oint\left(-\mu_0\,c^2\sqrt{1-\frac{v^2}{c^2}}+\frac{\mu_0\,c^2}{\sqrt{1-\frac{v^2}{c^2}}}+E_{\text{pot.}}\right)dt = \oint\left(\frac{\mu_0\,v^2}{\sqrt{1-\frac{v^2}{c^2}}}+E_{\text{pot.}}\right)dt = \oint\left(\sum_{k=1}^{3}\mu\,v_k\,\frac{dx_k}{dt}+E_{\text{pot}}\right)dt\,,$$

and upon partial integration over the closed periodic path, that will be:

$$= \oint\left(-\sum_{k=1}^{3}x_k\,\frac{d}{dt}\,\mu\,v_k+E_{\text{pot}}\right)dt = \oint\left(\sum_{k=1}^{3}x_k\,\frac{\partial E_{\text{pot}}}{\partial x_k}+E_{\text{pot}}\right)dt\,,$$

as a result of the equation of motion:

$$\frac{d}{dt}\,\mu\,v_k = -\frac{\partial}{\partial x_k}E_{\text{pot}}\,.$$

If one now makes the *assumption* that E_{pot} is a homogeneous function of degree – 1 in the x_k (**Coulomb** potential) then the integrand will vanish as a result of **Euler**’s theorem on homogeneous functions.] All that remains of (33′) is the theorem (33) that was to be proved.

However, with its help, it will follow from (29) that *when one carries the **Weyl** length ds around a spatially-closed quantum path, it will return to its initial value ds_0* when the still-undetermined factor K in (29) is taken to be:

$$K = -\frac{2i\,\pi\,\varepsilon}{h\,c}\,, \tag{34}$$

and α is stationary, i.e., it depends upon position, but not time. That result will refute the objection to **Weyl**’s theory (no. **54**, conclusion), in the event that one lets **Weyl**’s measure of length wander around only stationary **Bohr** quantum orbits, and the restriction of the allowed motion to such a situation is entirely natural within the context of **Bohr**’s theory.

56. Quantum-mechanical reinterpretation of Weyl's theory. – However, the **Bohr** orbits that are distinguished by quantum conditions are superseded by the ψ-states of **Schrödinger**'s theory, and the previous result will once more be called into question. Nonetheless, according to **F. London** (¹), **Weyl**'s theory can be extended in such a way that it is also appropriate to undulatory mechanics. Indeed, according to no. **52**, the state $\psi(x_1, x_2, x_3, x_4)$ corresponds to a current for which the coordinate increments δx_1, δx_2, δx_3, δx_4 will have a certain relationship to each other at every world-point. In no. **53**, a fifth coordinate x_5 with $\delta x_5 = c\ \delta t$ (τ = proper time) was introduced such that ψ would change along a five-dimensional streamline according to formula (24):

$$\psi = \psi_0 \exp\left[\frac{2i\pi}{h}\int\sum_5\left(-\frac{\varepsilon}{c}\Phi^k + \frac{h}{2i\pi}\frac{\partial \ln \alpha}{\partial x_k}\right)dx_k\right]. \tag{35}$$

Now, **London** also extended **Weyl**'s theory to five dimensions by associating the displacements δx_1 to δx_4 with a displacement $\delta x_5 = c\ d\tau$ and then replaced **Weyl**'s equation (26′) with:

$$\delta(ds^2) = ds^2 \cdot K \sum_{k=1}^{5} \varphi^k\ dx_k\ ,$$

so corresponding to (29):

$$ds^2 = (ds_0)^2 \cdot \exp\left[K\int\left(\sum_5 \Phi^k + \frac{1}{K}\frac{\partial \ln \alpha}{\partial x_k}\right)dx_k\right], \tag{36}$$

in which Φ^1 to Φ^5 are functions of x_1 to x_4 . As in **Weyl**'s theory, one will then get **Maxwell**'s equations precisely when one identifies Φ^1 to Φ^4 with the electromagnetic potentials but identifies Φ^4 with the **London** potential (16′). Namely, since the partial derivatives of Φ^1 to Φ^5 with respect to x_5 vanish, the components Rot^{ik5} in (31):

$$\mathrm{Rot}^{ikl}\, F = 0 \qquad (\text{in which } F^{ik} = \mathrm{Rot}^{ik}\,\Phi) \tag{37}$$

add nothing new to the **Maxwell** equations. If one now uses the special value for K in (34), namely, $K = -\,2i\ \pi\,\varepsilon\,/\,hc$, then the exponential functions in (35) and (36) will coincide, and with **London**, one will get:

$$\frac{\psi}{\psi_0} = \frac{ds^2}{ds_0^2}, \qquad \frac{\psi}{ds^2} = \frac{\psi_0}{ds_0^2}. \tag{38}$$

In the five-dimensional extension of **Weyl**'s theory, the magnitude of ds^2 will then change under translation of the gauge along a five-dimensional streamline by the same ratio as the ψ-function. The thing that behaves just like the **Weyl** measure will be the complex amplitude of the **Schrödinger** undulation as long as one lets the measure drift along the five-dimensional current.

(¹) **F. London**, Zeit. Phys. **42** (1927), pp. 375; Naturwissensch. Bd. **15**, Heft 8.

In particular, ds^2 will return to its initial value ds_0^2 when one traverses a *period* of ψ along the streamline.

57. Circular electron in the absence of a field. – The quantum mechanics of the relativistic electron that was developed up to now in this chapter started from the model of a charged mass-point. However, in the language of model building, according to **Uhlenbeck** and **Goudsmit**, the electron possesses a mechanical angular impulse $\frac{1}{2}\frac{h}{2\pi}$ that is coupled with a magnetic moment with the "anomalous" magnitude of:

$$1 \text{ magneton} = \frac{\varepsilon}{2\mu_0 c}\cdot\frac{h}{2\pi}, \tag{39}$$

in place of the "normally"-expected value of 1/2 magneton. Attempts to introduce that electron spin into wave mechanics go back to **Pauli** (^{1}), **Darwin** (^{2}), **Jordan** (^{3}), **Frenkel** (^{4}), **Ivanenko** and **Landau** (^{5}), and **Richter** (^{6}), among others, which arrived at the particular conclusion that the apparent electron spin was not an additional hypothesis that one was forced to add to the theory of the point electron after the fact, but was developed as a natural consequence of a simpler fundamental equation of wave mechanics. That requirement corresponded completely to the theory that **Dirac** (^{7}) had presented by means of an alteration of the previous fundamental equation (5):

$$\sum_k \left\{\left(p_k + \frac{\varepsilon}{c}\Phi_k\right)^2 + \mu_0^2 c^2, \psi\right\} = 0, \quad \text{where } p_k = \frac{h}{2i\pi}\frac{\partial}{\partial x_k}. \tag{40}$$

According to **Dirac**, one must deal with four different wave functions ψ_1, ψ_2, ψ_3, ψ_4 that satisfy the coupled system of equations:

$$\left\{ i\sum_k \gamma_k p_k + \mu_0 c, \psi_\zeta \right\} = 0 \qquad \text{for} \qquad \zeta = 1, 2, 3, 4, \tag{41}$$

which is *linear* in the p_k, *in the absence of a field.* The coefficients γ_k in that are therefore *operators* that do not include the p and q, so they will commute with them. However, if γ_k is applied to ψ_ζ then the result $\{\gamma_k, \psi_\zeta\}$ will be defined by the series:

(^{1}) **W. Pauli**, Zeit. Phys. **43** (1927), pp. 601.
(^{2}) **C. G. Darwin**, Proc. Roy. Soc. **116** (1927), pp. 227.
(^{3}) **P. Jordan**, Zeit. Phys. **44** (1927), pp. 1.
(^{4}) **J. Frenkel**, **47** (1928), pp. 786.
(^{5}) **D. Ivanenko** and **L. Landau**, Zeit. Phys. **48** (1928), pp. 340.
(^{6}) **C. F. Richter**, Proc. Nat. Acad. America **13** (1927), pp. 476.
(^{7}) **P. A. M. Dirac**, Proc. Roy. Soc. **117** (1928), pp. 610.

$$\{\gamma_k,\ \psi_\zeta\} = \sum_{\zeta} \gamma_k^{\zeta\zeta'} \cdot \psi_{\zeta'}\,, \tag{42}$$

which has constant coefficients $\gamma_k^{\zeta\zeta'}$ whose values will be given soon. Those coefficients shall be regarded as matrix components of the operator γ_k, which corresponds to the equation that arises from (42) upon multiplying by $\psi_{\zeta''}$ and integrating over space:

$$\int \psi_{\zeta''}\,\gamma_k\,\psi_\zeta\,dv = \sum_{\zeta'}\gamma_k^{\zeta\zeta'}\int \psi_{\zeta''}\,\psi_{\zeta'}\,dv = \gamma_k^{\zeta\zeta''}, \tag{43}$$

in which the ψ are assumed to be mutually orthogonal and normalized to 1.

Since ψ includes the index $\zeta = 1, 2, 3, 4$ here, the number of ψ-solutions quadrupled in comparison to the family of solutions that was previously considered. However, that will be reduced to one-half that much (cf., *infra*).

Now, **Dirac** demanded that the operators γ_k should have a form that *in the absence of a field*, the previous wave equation of the point electron *in the absence of a field* will emerge from the Ansatz (41):

$$\left\{\sum p_k^2 + \mu_0^2\,c^2, \psi_\zeta\right\} = 0\,. \tag{44}$$

If one now applies the operation to (41) then one will get the equation:

$$\begin{aligned} 0 &= \left\{\left(-i\sum_l \gamma_l\,p_l + \mu_0\,c\right)\cdot\left(i\sum_k \gamma_k\,p_k + \mu_0\,c\right), \psi_\zeta\right\} \\ &= \left\{\sum_k \gamma_k^2\,p_k^2 + \sum_{k<l}\sum (\gamma_k\,\gamma_l + \gamma_l\,\gamma_k + p_k\,p_l + \mu_0^2\,c^2, \psi_\zeta\right\}, \end{aligned}$$

and that will be the previous equation (44) when one demands of the operators γ that:

$$\gamma_k^2 = 1, \qquad \gamma_k\,\gamma_l + \gamma_l\,\gamma_k = 0 \qquad \text{for} \qquad k \neq l. \tag{45}$$

That requirement will be fulfilled when γ_k are defined by the following matrix components $\gamma_k^{\zeta\zeta'}$:

$$\left.\begin{aligned}
\gamma_1 &= \begin{Bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{Bmatrix}, \quad \gamma_2 = \begin{Bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{Bmatrix}, \\
\gamma_3 &= \begin{Bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{Bmatrix}, \quad \gamma_4 = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}.
\end{aligned}\right\} \tag{46}$$

In the absence of a field, **Dirac**'s Ansatz will then give nothing new about the theory of point electron (but it probably will in the presence of a field, cf., no. **58**). The rules for γ in (45) can also be summarized in the form:

$$\gamma_k\,\gamma_l + \gamma_l\,\gamma_k = 2 \cdot \delta_{kl} \qquad (\delta_{kl} = 1 \text{ for } k = l \text{ and } = 0 \text{ for } k \neq l). \tag{47}$$

We shall now show the invariance of the **Dirac** wave equation under ***Lorentz*** *transformations* of the coordinates. One will obtain the form four-vector p from the four linear equations (rotation in x_1, x_2, x_3, x_4-space):

$$p'_k = \sum_l a_{kl}\, p_l \qquad (k, l = 1, 2, 3, 4) \tag{48}$$

with the orthogonal numerical coefficients a :

$$\sum_k a_{kl}\, a_{km} = \delta_{lm}\,, \qquad \sum_m a_{km}\, a_{lm} = \delta_{kl}\,. \tag{48′}$$

Introducing that into the wave equation (41) will then give:

$$\left\{i \sum \gamma'_n\, p'_n + \mu_0\, c, \psi\right\} = 0\,, \tag{49}$$

in which:

$$\gamma'_n = \sum_m a_{nm}\, \gamma_m\,, \tag{49′}$$

i.e., γ_1 to γ_4 transform like a four-vector. From (49′), one will have:

$$\gamma'_k\,\gamma'_l + \gamma'_l\,\gamma'_k = \sum_m \sum_n a_{km}\, a_{ln} (\gamma_m\,\gamma_n + \gamma_n\,\gamma_m) = 2 \sum_m a_{km}\, a_{lm} = 2\, \delta_{kl}\,, \tag{50}$$

such that the γ'_k in the new coordinate system obey the same relations (37) as the γ_k did in the original coordinate system.

Dirac further showed that the Lorentz transformation is, at the same time, a *canonical transformation* that is always allowed quantum-mechanically.

Instead of employing the γ in (46), one can also take other γ' as a starting point for fixed coordinates, which arise from the ones above by an arbitrary orthogonal transformation (49′). They do not need to be **Minkowskian** (1, 2, 3 real, 4 imaginary) to begin with. For example, instead of (46), one can start with the γ-matrices:

$$\left.\begin{array}{ll}
\gamma_1 = \left\{\begin{matrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{matrix}\right\}, & \gamma_2 = \left\{\begin{matrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}\right\}, \\
\gamma_3 = \left\{\begin{matrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{matrix}\right\}, & \gamma_4 = \left\{\begin{matrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{matrix}\right\},
\end{array}\right\} \tag{50′}$$

without changing any of the physical results.

In full generality, one can also employ:

$$\gamma'_n = \Lambda^{-1}\,\gamma_n\,\Lambda$$

as the starting matrices, instead of the γ_n , where Λ is an arbitrary four-dimensional matrix, and Λ^{-1} is its reciprocal:

$$(\gamma'_n)^{jk} = \sum_s\sum_t (\Lambda^{-1})^{js}(\gamma_n)^{st}(\Lambda)^{tk} .$$

58. Circular electron in a field. – The transition to arbitrary electromagnetic fields will be completed when one replaces the impulse components p_k with $p_k + (e / c)\,\Phi_k$ with the four-potential Φ_k . **Dirac** then obtained the fundamental linear equation of the electron:

$$\left\{ i\sum_k \gamma_k \left(p_k + \frac{\varepsilon}{c}\Phi_k \right) + \mu_0\, c, \psi_\zeta \right\} = 0 \tag{51}$$

as a generalization of (41). If one then multiplies that by the conjugate operator, so one defines:

$$\left\{ \left[-i\sum_l \gamma_l \left(p_l + \frac{\varepsilon}{c}\Phi_l \right) + \mu_0\, c \right] \left[i\sum_k \gamma_k \left(p_k + \frac{\varepsilon}{c}\Phi_k \right) + \mu_0\, c \right], \psi_\zeta \right\} = 0 ,$$

then upon multiplying things out, one will get:

$$\left\{\mu_0^2 c^2 + \sum_k \gamma_k^2 \left(p_k + \frac{\varepsilon}{c}\Phi_k\right)^2 + \sum_k \sum_l \gamma_k \gamma_l \left(p_k + \frac{\varepsilon}{c}\Phi_k\right)\left(p_l + \frac{\varepsilon}{c}\Phi_l\right), \psi_\zeta\right\} = 0 . \qquad (51')$$

Since $\gamma_k^2 = 1$, the first two terms are identical to the previous wave equation (40) for the electron in the field. However, since $\gamma_k \gamma_l = - \gamma_l \gamma_k$ for $k \neq l$, they must be combined with the term:

$$\left\{\gamma_1 \gamma_2 \left[\left(p_1 + \frac{\varepsilon}{c}\Phi_1\right)\left(p_2 + \frac{\varepsilon}{c}\Phi_2\right) - \left(p_2 + \frac{\varepsilon}{c}\Phi_2\right)\left(p_1 + \frac{\varepsilon}{c}\Phi_1\right)\right] + \cdots, \psi_\zeta\right\}$$

$$= \left\{\frac{\varepsilon}{c}\gamma_1 \gamma_2 (p_1 \Phi_2 - \Phi_2 \, p_1 - p_2 \, \Phi_1 + \Phi_1 \, p_2) + \cdots, \psi_\zeta\right\}$$

$$= \frac{\varepsilon}{c}\frac{h}{2i\pi}\gamma_1 \gamma_2 \psi_\zeta \cdot \left(\frac{\partial \Phi_2}{\partial x_1} - \frac{\partial \Phi_1}{\partial x_2}\right) + \cdots = \frac{\varepsilon}{c}\frac{h}{2i\pi}\mathrm{Rot}_{12}\,\Phi \cdot \gamma_1 \gamma_2 \psi_\zeta + \cdots$$

With the help of the abbreviations:

$$\mathrm{Rot}_{kl}\,\Phi = \mathfrak{F}_{kl} = - \mathfrak{F}_{lk} \quad \text{and} \qquad \frac{h}{2i\pi}\cdot \gamma_k \gamma_l = \mathfrak{G}_{kl} = - \mathfrak{G}_{lk} , \qquad (52)$$

one will ultimately obtain the total wave equation in the form:

$$\left\{\frac{1}{2}\mu_0^2 c^2 + \frac{1}{2\mu_0}\sum_k \gamma_k^2 \left(p_k + \frac{\varepsilon}{c}\Phi_k\right)^2 + \frac{\varepsilon}{2\mu_0 c}(\mathfrak{F}, \mathfrak{G}), \psi_\zeta\right\} = 0 , \qquad (53)$$

with the scalar product $(\mathfrak{F}, \mathfrak{G})$, which consists of six terms. If one introduces the notations:

$$\left.\begin{array}{llllll} \mathfrak{F}_{23} = \mathfrak{H}_x, & \mathfrak{F}_{31} = \mathfrak{H}_y, & \mathfrak{F}_{12} = \mathfrak{H}_z, & \mathfrak{F}_{14} = -i\,\mathfrak{E}_x, & \mathfrak{F}_{24} = -i\,\mathfrak{E}_y, & \mathfrak{F}_{34} = -i\,\mathfrak{E}_z, \\ \mathfrak{G}_{23} = \mathfrak{Q}_x, & \mathfrak{G}_{31} = \mathfrak{Q}_y, & \mathfrak{G}_{12} = \mathfrak{Q}_z, & \mathfrak{G}_{14} = +i\,\mathfrak{P}_x, & \mathfrak{G}_{24} = +i\,\mathfrak{P}_y, & \mathfrak{G}_{34} = +i\,\mathfrak{P}_z \end{array}\right\} \qquad (54)$$

then one will get the supplementary term in the form:

$$\frac{\varepsilon}{2\mu_0 c}(\mathfrak{F}, \mathfrak{G}) = \frac{\varepsilon}{2\mu_0 c}(\mathfrak{H}, \mathfrak{Q}) + \frac{\varepsilon}{2\mu_0 c}(\mathfrak{E}, \mathfrak{P}) . \qquad (54')$$

One can then think of the wave equation as something that arises from a classical energy function with supplementary terms that suggest that the electron possesses a supplementary magnetic energy of $\frac{\varepsilon}{2\mu_0\, c}(\mathfrak{H},\mathfrak{Q})$ and a supplementary electric energy of $\frac{\varepsilon}{2\mu_0\, c}(\mathfrak{E},\mathfrak{P})$ that originate in a proper magnetic moment of $\mathfrak{q} = \frac{\varepsilon}{2\mu_0\, c}\mathfrak{Q}$ and a proper electric moment of $\mathfrak{p} = \frac{\varepsilon}{2\mu_0\, c}\mathfrak{P}$, which is *imaginary*, as one sees from (56′). Wave-mechanically, $\mathfrak{Q}$ and $\mathfrak{P}$ ($\mathfrak{G}$, resp.) are then the operators that were given in (54) and act on ψ in sense of the defining equation:

$$\{\mathfrak{G}_{kl},\ \psi_\zeta\} = \sum_{\zeta} \mathfrak{G}_{kl}^{\zeta\zeta'} \cdot \psi_{\zeta'}\,. \tag{55}$$

In detail, when one employs (46) and (52), one can then derive:

$$\mathfrak{Q}_x = \frac{h}{2\pi}\begin{Bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{Bmatrix}, \quad \mathfrak{Q}_y = \frac{h}{2\pi}\begin{Bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{Bmatrix}, \quad \mathfrak{Q}_z = \frac{h}{2\pi}\begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}, \tag{56}$$

$$i\,\mathfrak{P}_x = \frac{h}{2\pi}\begin{Bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{Bmatrix}, \quad i\,\mathfrak{P}_y = \frac{h}{2\pi}\begin{Bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{Bmatrix}, \quad i\,\mathfrak{P}_z = \frac{h}{2\pi}\begin{Bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{Bmatrix}, \tag{56′}$$

and one will have:

$$\mathfrak{Q}_x^2 = \mathfrak{Q}_y^2 = \mathfrak{Q}_z^2 = \left(\frac{h}{2\pi}\right)^2 \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix} = -\mathfrak{P}_x^2 = -\mathfrak{P}_y^2 = -\mathfrak{P}_z^2\,. \tag{57}$$

In order to derive the *energy term* for an electron in a central electric field, **Dirac** went over to polar coordinates by a canonical transformation and then found solutions that would coincide, in the first approximation, with the results that were also confirmed by experiment that **Darwin** (^{1}) had achieved from the incomplete wave equation (40) by adding a suitable term that represented the electron spin. The great advance of the **Dirac** equation lies in the fact that it accounted for those phenomena in an uncontrived way and with no supplementary hypotheses.

(^{1}) **C. G. Darwin**, Proc. Roy. Soc. **115** (1927), pp. 1.

Sommerfeld's relativistic fine structure of *H*-atom terms, which **Goudsmit** regarded as an effect of a top impulse, was derived by **Gordon** (^{1}) on the basis of **Dirac**'s theory. The Zeeman effect and the line intensities were derived by **Dirac** himself (^{2}).

The necessity of introducing an imaginary electric moment $\mathfrak{P}$, along with the real magnetic moment Q [Q^2 is a positive operator, while $\mathfrak{P}^2$ is a negative one, cf., (57)] in order to obtain a relativistically-invariant theory was first recognized and employed by **Frenkel** (^{3}).

The necessity of completing the transition from point to top electrons is the reason for why we must forgo reporting upon the stimulating "Untersuchungen zum Problem der Quantenelektrik" by **G. Mie** (^{4}).

(^{1}) **W. Gordon**, Zeit. Phys. **48** (1928), pp. 11.
(^{2}) **P. A. M. Dirac**, Proc. Roy. Soc. **118** (1928), pp. 351.
(^{3}) **J. Frenkel**, Zeit. Phys. **47** (1928), pp. 786.
(^{4}) **G. Mie**, Ann. Phys. (Leipzig) **85** (1928), pp. 711.

CHAPTER VII

QUANTUM ALGEBRA AND TRANSFORMATIONS

59. Correspondence. – The coupling between the classical and undulatory mechanics for a system with the coordinates $q_1, \ldots, q_N$ and canonically-conjugate impulses $p_1, \ldots, p_N$ comes about by a reinterpretation of the impulse in terms of operators:

$$p_K = \frac{h}{2i\pi}\frac{\partial}{\partial q_K}, \qquad -E = \frac{h}{2i\pi}\frac{\partial}{\partial t}.$$

In that way, the **Hamilton** equation:

$$H(q, p) = E = 0$$

will then become the **Schrödinger** partial differential equation:

$$\left\{H\left(q, \frac{h}{2\pi i}\frac{\partial}{\partial q}\right) + \frac{h}{2\pi i}\frac{\partial}{\partial t}, \psi(q,t)\right\} = 0 \cdot \psi(q, t)\,.$$

That reinterpretation might now be suggested by using boldface for the $\boldsymbol{q}$ and $\boldsymbol{p} = \frac{h}{2i\pi}\frac{\partial}{\partial q}$ and functions of them. The equation:

$$\boldsymbol{H}(\boldsymbol{q}, \boldsymbol{p}) - \boldsymbol{E} = 0$$

shall then mean precisely the same thing as the foregoing equation, namely, when the operator $H\left(q, \frac{h}{2\pi i}\frac{\partial}{\partial q}\right) + \frac{h}{2\pi i}\frac{\partial}{\partial t}$ is applied to a function $\psi(q, t)$, that will be equal to the 0 operator applied to ψ, i.e., $0 \cdot \psi$. One then has the *commutation rules* for the $\boldsymbol{p}$ and $\boldsymbol{q}$:

$$\frac{h}{2i\pi}\frac{\partial}{\partial q_K} q_K \psi - q_K \frac{h}{2i\pi}\frac{\partial}{\partial q_K}\psi = \frac{h}{2i\pi}\psi\,, \quad \text{etc.,}$$

i.e.:

$$\left.\begin{array}{ll} \boldsymbol{p}_K\,\boldsymbol{q}_K - \boldsymbol{q}_K\,\boldsymbol{p}_K = \dfrac{h}{2\pi i}, & \boldsymbol{p}_K\,\boldsymbol{q}_L - \boldsymbol{q}_L\,\boldsymbol{p}_K = 0 \quad \text{for } L \neq K, \\ \boldsymbol{p}_K\,\boldsymbol{p}_L - \boldsymbol{p}_L\,\boldsymbol{p}_K = 0, & \boldsymbol{p}_K\,\boldsymbol{q}_L - \boldsymbol{q}_L\,\boldsymbol{p}_K = 0. \end{array}\right\} \tag{1}$$

Moreover, the usual rules of calculation concerning association and distribution are still true, but commutation is allowed only for addition of $\boldsymbol{p}$ and $\boldsymbol{q}$, but not multiplication (^{1}). One can then calculate with the p and q just as one does with ordinary numbers, except that one must observe the commutation rules (1), which must also subsume the coordinate $\boldsymbol{t}$ and the conjugate impulse $-\boldsymbol{E}$. In that way, one arrives at an algebra of the quantities $\boldsymbol{q}$ and $\boldsymbol{p}$. It is the foundation for *quantum mechanics*, as it was already presented *before* **Schrödinger**'s theory by **Heisenberg**, **Born**, **Jordan** (^{2}) and **Dirac** (^{3}) as a method obtaining a numerical relationship between observable quantities that would reproduce classical mechanics only incompletely. In that way, the exhibiting of those commutation rules was the end result of a systematic comparison of the classical-mechanical expectation that quantum-mechanical reality should be a quantitative sharpening of ***Bohr****'s correspondence principle*. We would like to pursue the correspondence between classical mechanics and the quantum mechanics that is built upon quantum algebra in more detail.

The *correspondence* is exhibited by the association:

$$[x, y] = \sum_L \left(\frac{\partial x}{\partial p_L} \frac{\partial y}{\partial q_L} - \frac{\partial y}{\partial p_L} \frac{\partial x}{\partial q_L} \right) \qquad \textit{classically}, \tag{A}$$

$$[\boldsymbol{x}, \boldsymbol{y}] = (\boldsymbol{x}\,\boldsymbol{y} - \boldsymbol{y}\,\boldsymbol{x}) \qquad \textit{quantum-theoretical}, \tag{B}$$

which assigns two meaning to the **Poisson** bracket $[x\,(q, p), y\,(q, p)]$, namely, a classical one and a quantum-theoretical. In (A), x and y mean arbitrary functions of q and p. In (B), $\boldsymbol{x}$ and $\boldsymbol{y}$ mean the same functions of $\boldsymbol{q}$ and $\boldsymbol{p}$, except that the sequence of the factors x and y in the product must be symmetrized, so in place of $\boldsymbol{x} = \boldsymbol{p}\,\boldsymbol{q}$, one shall have, e.g., $\boldsymbol{x} = \frac{1}{2}(\boldsymbol{p}\,\boldsymbol{q} + \boldsymbol{q}\,\boldsymbol{p})$, in order for the commutability to also be ensured by the quantum-algebraic function.

If one sets $x\,(q, p) = p_K$, $y\,(q, p) = q_L$, in particular, then one will get:

$$[p_K, q_L] = 1 \quad \text{for } K = L, \qquad = 0 \quad \text{for } K \neq L .$$

The correspondence (A) → (B) will then lead to the demand that:

$$[\boldsymbol{p}_K, \boldsymbol{q}_L] = \mathbf{1} \quad \text{for } K = L, \qquad = 0 \quad \text{for } K \neq L . \tag{2}$$

One will be correspondingly led to the demands that:

$$[\boldsymbol{p}_K, \boldsymbol{p}_L] = \mathbf{0} , \quad [\boldsymbol{q}_K, \boldsymbol{q}_L] = \mathbf{0} . \tag{2′}$$

(^{1}) Cf., no. **44**.

(^{2}) **W. Heisenberg**, Zeit. Phys. **33** (1925), pp. 33; **M. Born** and **P. Jordan**, *ibidem*, **34** (1925), pp. 858; *ibid.*, **35** (1926), pp. 557.

(^{3}) **P. A. M. Dirac**, Proc. Roy. Soc. **109** (1925), pp. 642; *ibid.*, **110** (1926), pp. 561; *ibid.*, **111** (1926), pp. 281; *ibid.*, **112** (1926), pp. 611.

However, (2) and (2′) are identical to *commutation rules* (1) that were cited above, due to the meaning (B) of the **Poisson** bracket in quantum theory. That was the path along which the commutation rules were first found.

If one sets x and y on the right-hand side of (A) equal to $x\,(p, q)$ and q_K in one case and equal to p_K and $y\,(q, p)$ another time then one will get the following special equation from (A):

$$\left.\begin{aligned} \frac{\partial x(q,p)}{\partial p_K} &= [x, q_K], \\ \frac{\partial y(q,p)}{\partial p_K} &= [p_K, y]. \end{aligned}\right\} \tag{3}$$

The correspondence (A) → (B) then requires that the quantum-algebraic differential quotients must be defined as follows:

$$\left.\begin{aligned} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{p}_K} &= [\boldsymbol{x}, \boldsymbol{q}_K] = \frac{2i\pi}{h}(\boldsymbol{x}\,\boldsymbol{q}_K - \boldsymbol{q}_K\,\boldsymbol{x}), \\ \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{q}_K} &= [\boldsymbol{p}_K, \boldsymbol{y}] = \frac{2i\pi}{h}(\boldsymbol{p}_K\,\boldsymbol{y} - \boldsymbol{y}\,\boldsymbol{p}_K). \end{aligned}\right\} \tag{4}$$

In order to learn about the quantum operation that corresponds to the classical differential quotient with respect to time t, we consider the canonical equations of motion in classical mechanics:

$$\dot{q}_K = \frac{\partial H(q,p)}{\partial p_K}, \qquad \dot{p}_K = -\frac{\partial H(q,p)}{\partial q_K} \qquad (K = 1, 2, \dots, N), \tag{5}$$

take an arbitrary function $z\,(q, p)$, and define:

$$\dot{z} = \sum_L \left(\frac{\partial z}{\partial q_L}\dot{q}_L + \frac{\partial z}{\partial p_L}\dot{p}_L \right) = \sum_L \left(\frac{\partial z}{\partial q_L}\frac{\partial H}{\partial p_L} - \frac{\partial z}{\partial p_L}\frac{\partial H}{\partial q_L} \right) = [H, z] \tag{5′}$$

as an expression that summarizes the canonical equations of motion and subsumes (5) as the special cases $z = q_K$ and $z = p_K$. We now employ the correspondence of the **Poisson** brackets in order to quantum-theoretically define:

Let the dotted form $\dot{\boldsymbol{z}}$ of a q-function $\boldsymbol{z}\,(\boldsymbol{q}, \boldsymbol{p})$ *with respect to a certain **Hamiltonian** function* $\boldsymbol{H}\,(\boldsymbol{q}, \boldsymbol{p})$ be the function:

$$\dot{\boldsymbol{z}} = [\boldsymbol{H}, \boldsymbol{z}] = \frac{2i\pi}{h}(\boldsymbol{H}\,\boldsymbol{z} - \boldsymbol{z}\,\boldsymbol{H}). \tag{6}$$

$\dot{\boldsymbol{z}}$ corresponds to the time derivative $\dot{z}$ of the classical theory, and for $\boldsymbol{z} = \boldsymbol{q}_K$ ($\boldsymbol{z} = \boldsymbol{p}_K$, resp.), (6) will correspond to the classical *canonical equations of motion:*

$$\dot{\boldsymbol{q}}_K = \frac{2i\pi}{h}(\boldsymbol{H}\,\boldsymbol{q}_K - \boldsymbol{q}_K\,\boldsymbol{H}), \qquad \dot{\boldsymbol{p}}_K = \frac{2i\pi}{h}(\boldsymbol{p}_K\,\boldsymbol{H} - \boldsymbol{H}\,\boldsymbol{p}_K). \tag{6′}$$

For the special case $\boldsymbol{z}(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{H}(\boldsymbol{q}, \boldsymbol{p})$, it will further follow from (6) that:

$$\dot{\boldsymbol{H}} = [\boldsymbol{H}, \boldsymbol{H}] = 0. \tag{6″}$$

The dotted form of the **Hamiltonian** function vanishes. That is the quantum-mechanical counterpart to the *law of conservation of energy* $\dot{H} = 0$ in classical mechanics.

60. Canonical transformations. Angle variables (1). – In order to solve the quantum-mechanical problem, it is usually necessary to go from the original variables q_k, p_k to the new variables Q_k, P_k by a *transformation:*

$$\boldsymbol{Q}_K = \boldsymbol{Q}_K(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N), \qquad \boldsymbol{P}_K = \boldsymbol{P}_K(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N).$$

The demand that the new variables, like the old ones, are once more *canonical*, i.e., they must satisfy the commutation rules:

$$\left.\begin{array}{l} [\boldsymbol{P}_K, \boldsymbol{Q}_L] = 1 \quad \text{for } K = L, \quad \text{for } K \neq L, \\ [\boldsymbol{P}_K, \boldsymbol{P}_L] = 0, \quad [\boldsymbol{Q}_K, \boldsymbol{Q}_L] = 0, \end{array}\right\} \tag{7}$$

will be fulfilled by any *transformation of the form:*

$$\boldsymbol{Q}_K = \boldsymbol{T}^{-1}(\boldsymbol{q}, \boldsymbol{p}) \cdot \boldsymbol{q}_K \cdot \boldsymbol{T}(\boldsymbol{q}, \boldsymbol{p}), \qquad \boldsymbol{P}_K = \boldsymbol{T}^{-1} \cdot \boldsymbol{p}_K \cdot \boldsymbol{T}, \tag{8}$$

in which $\boldsymbol{T}(\boldsymbol{q}, \boldsymbol{p})$ is an arbitrary function of the $\boldsymbol{q}$, $\boldsymbol{p}$, and $\boldsymbol{T}^{-1}$ means its reciprocal, which is defined by $\boldsymbol{T}^{-1} \cdot \boldsymbol{T} = \boldsymbol{T} \cdot \boldsymbol{T}^{-1} = \boldsymbol{1}$. Upon applying it to the sums and products of the $\boldsymbol{p}$ and $\boldsymbol{q}$, one will easily arrive at the more general formula:

$$\boldsymbol{F}(\boldsymbol{Q}, \boldsymbol{P}) = \boldsymbol{F}(\boldsymbol{T}^{-1}\boldsymbol{q}\,\boldsymbol{T}, \boldsymbol{T}^{-1}\boldsymbol{p}\,\boldsymbol{T}) = \boldsymbol{T}^{-1}\boldsymbol{F}(\boldsymbol{q}, \boldsymbol{p})\,\boldsymbol{T}. \tag{8′}$$

Let a mechanical system be *classically* characterized by a certain **Hamiltonian** function $F(q, p)$ and quantum-mechanically by a symmetrized **Hamiltonian** function $\boldsymbol{H}(\boldsymbol{q}, \boldsymbol{p})$.

One might now ask about the quantum-theoretical *energy values* (spectral terms), and later about the frequencies, intensities, and polarizations of the light that is emitted by the system. The solution will be achieved in the following way: One goes from the coordinates q and impulses p to the new coordinates w (viz., the angle variables) and impulses J (action variables) by a canonical transformation:

---

(1) **M. Born**, **W. Heisenberg**, and **P. Jordan**, Zeit. Phys. **35** (1926), pp. 667; **P. Jordan**, *ibidem*, **37** (1926), pp. 383; **G. Wentzel**, *ibidem*, **37** (1926), pp. 80.

$$\boldsymbol{w}_K = \boldsymbol{T}^{-1}(\boldsymbol{q}, \boldsymbol{p}) \cdot \boldsymbol{q}_K \cdot \boldsymbol{T}, \qquad \boldsymbol{J}_K = \boldsymbol{T}^{-1} \cdot \boldsymbol{p}_K \cdot \boldsymbol{T}, \qquad \text{so} \qquad \boldsymbol{F}(\boldsymbol{w}, \boldsymbol{J}) = \boldsymbol{T}^{-1}\boldsymbol{F}(\boldsymbol{q}, \boldsymbol{p})\,\boldsymbol{T}$$

in which the transforming function ***T* (*q*, *p*)** is chosen such that in addition to the canonical character of the ***w*** and ***J***, which is ensured by (9) with no further conditions, namely:

$$\left.\begin{array}{l} [\boldsymbol{J}_K, \boldsymbol{w}_L] = 1 \quad \text{for } K = L, \quad \text{for } K \neq L, \\ [\boldsymbol{J}_K, \boldsymbol{J}_L] = 0, \quad [\boldsymbol{w}_K, \boldsymbol{w}_L] = 0, \end{array}\right\} \tag{10}$$

the following conditions are also fulfilled:

a) If one replaces the ***q*** and ***p*** in the **Hamiltonian** function $\boldsymbol{H}(\boldsymbol{q}, \boldsymbol{p})$ with the functions $\boldsymbol{q}_K(\boldsymbol{w}, \boldsymbol{J})$, $\boldsymbol{p}_K(\boldsymbol{w}, \boldsymbol{J})$ that follow by inversion of (9) then the function that thus arises $\boldsymbol{H}(\boldsymbol{q}(\boldsymbol{w}, \boldsymbol{J}), \boldsymbol{p}(\boldsymbol{w}, \boldsymbol{J})) = \boldsymbol{H}^*(\boldsymbol{w}, \boldsymbol{J})$ will include only the ***J***, but not the ***w*** :

$$\boldsymbol{H}(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{H}^*(\boldsymbol{J}).$$

b) The functions $\boldsymbol{q}_K(\boldsymbol{w}, \boldsymbol{J})$, $\boldsymbol{p}_K(\boldsymbol{w}, \boldsymbol{J})$ shall have the following forms as a series:

$$\boldsymbol{q}_K = \sum_\tau \boldsymbol{q}_K^\tau(\boldsymbol{J}) \cdot e^{i(\tau w)}, \qquad \boldsymbol{p}_K = \sum_\tau \boldsymbol{p}_K^\tau(\boldsymbol{J}) \cdot e^{i(\tau w)}. \tag{11}$$

The $(\tau\,\boldsymbol{w})$ in that describes an abbreviation for $\tau_1\,\boldsymbol{w}_1 + \ldots + \tau_N\,\boldsymbol{w}_N$, and the associated coefficient $\boldsymbol{q}_K^\tau(\boldsymbol{J})$ would be denoted more precisely by $\boldsymbol{q}_K^{\tau_1 \cdots \tau_N}(\boldsymbol{J}_1, \ldots, \boldsymbol{J}_N)$, and $\sum_\tau$ is written as an abbreviation for $\sum_{\tau_1} \cdots \sum_{\tau_N}$, where each τ_K is summed over all whole numbers from $-\infty$ to $+\infty$. The p_K and q_K are then periodic in the w with period 2π.

If the condition b) is fulfilled then it will follow that any function $\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{q})$ that is composed from the ***p*** and ***q*** by any combination of addition and multiplication can be represented in the corresponding form of a series:

$$\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{q}) = \sum_\tau \boldsymbol{x}^\tau(\boldsymbol{J}) \cdot e^{i(\tau w)}. \tag{11′}$$

Due to (8) in no. **44**, that is identical to:

$$\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{q}) = \sum_\tau e^{i(\tau w)}\, \boldsymbol{x}^\tau\left(\boldsymbol{J} + \frac{\tau h}{2\pi}\right). \tag{11″}$$

$\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{q})$ will also be periodic of period 2π in the ***w*** then.

Now let the transformation to angle and action variables ***w*** and ***J*** be effected by a suitably-chosen transforming ***T***, so the transformed **Hamiltonian** function is found to be $\boldsymbol{H}^*(\boldsymbol{J})$. That will then give rise to the **Schrödinger** equation:

$$\left\{H^*\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right), \psi^*(w)\right\} = \{E \cdot \psi^*(w)\}. \tag{12}$$

If the $H^*(J)$ in that is developed in the power series:

$$H^*(J) = \sum_\tau A_{\tau_1 \cdots \tau_N} \cdot (J_1)^{\tau_1} \cdot (J_2)^{\tau_2} \cdots (J_N)^{\tau_N} \tag{12$'$}$$

then the oscillation equation $\left(J_K = \frac{h}{2i\pi} \cdot \frac{\partial}{\partial w_K}\right)$:

$$\sum_\tau A_{\tau_1 \cdots \tau_N} \left(\frac{h}{2i\pi}\right)^{\tau_1+\tau_2+\cdots+\tau_N} \frac{\partial^{\tau_1}}{\partial w_1^{\tau_1}} \frac{\partial^{\tau_2}}{\partial w_2^{\tau_2}} \cdots \frac{\partial^{\tau_N}}{\partial w_N^{\tau_N}} \psi^*(w_1, \ldots, w_N) = E \cdot \psi^*(w_1, \ldots, w_N) \tag{13}$$

will be solved by the Ansatz $\psi^*(w) = B e^{i(nw)}$, which is:

$$\psi^*(w_1, \ldots, w_N) = C \cdot e^{i(n_1 w_1 + \cdots + n_N w_N)} \tag{14}$$

when written out in detail, and in which $n_1, \ldots, n_N$ are arbitrary numerical values. If one chooses the latter to be *whole numbers* then one will succeed in making $\psi^*(w)$ *periodic* in the w with period 2π. When the Ansatz (14) is substituted in (13), that will give the eigenvalue that belongs to (14):

$$\sum_\tau A_{\tau_1 \cdots \tau_N} \left(\frac{h}{2i\pi}\right)^{\tau_1+\tau_2+\cdots+\tau_N} \cdot (n_1 i)^{\tau_1} \cdot (n_2 i)^{\tau_2} \cdots \cdot (n_N i)^{\tau_N} = E, \tag{15}$$

which one can denote more precisely by $E_{n_1 \cdots n_N}$. One can normalize the still-undetermined factor C of ψ such that one has:

$$\int_0^{2\pi} \cdots \int_0^{2\pi} \psi \tilde{\psi} \cdot dw_1 \cdots dw_N = 1,$$

i.e.:

$$C = (2\pi)^{-N/2}.$$

The eigenvalue problem will then be solved when one succeeds in turning the Hamiltonian function into a function $\boldsymbol{H}^*(\boldsymbol{J}_1,\dots,\boldsymbol{J}_N)$ of the action variables alone by a canonical transformation $(\boldsymbol{q}, \boldsymbol{p}) \rightarrow (\boldsymbol{w}, \boldsymbol{J})$. The eigenfunctions will then be the *exponential functions* (14) with whole-number $n_1, \dots, n_N$, and from (15), (12′), the eigenvalue will possess the simple form:

$$E = H^*\left(\frac{n_1 h}{2\pi}, \frac{n_2 h}{2\pi}, \dots, \frac{n_N h}{2\pi}\right). \tag{15′}$$

One can subsequently express that result without speaking of **Schrödinger**'s undulatory mechanics as a result of quantum algebra in the following way:

If one succeeds in going from the **Hamiltonian** function $\boldsymbol{H}(\boldsymbol{p}, \boldsymbol{q})$ to angle and action variables $\boldsymbol{w}, \boldsymbol{J}$ by a canonical transformation $\boldsymbol{T}$ such that $\boldsymbol{H}$ will go to a function $\boldsymbol{H}^*(\boldsymbol{J}_1,\dots,\boldsymbol{J}_N)$ then one will get the quantum-theoretic energy values when one replaces the arguments $\boldsymbol{J}_1, \dots, \boldsymbol{J}_N$ in the latter function with the quantities $n_1 h / 2\pi, \dots, n_N h / 2\pi$ with whole-number $n_1, \dots, n_N$.

61. Canonical transformations (continuation). – In classical mechanics, one will get a canonical transformation of the q, p into new variables Q, P that does not change the form of **Hamilton**'s equations of motion, when one poses the equations:

$$q_k = \frac{\partial S}{\partial p_k}, \qquad P_k = \frac{\partial S}{\partial Q_k} \qquad [S(Q, p)]$$

with the help of an arbitrary "action function," solves them for $Q_k = Q_k(q, p)$, $P_k = P_k(q, p)$.

In quantum mechanics, as well:

$$\boldsymbol{q}_k = \frac{\partial \boldsymbol{S}}{\partial \boldsymbol{p}_k}, \qquad \boldsymbol{P}_k = \frac{\partial \boldsymbol{S}}{\partial \boldsymbol{Q}_k} \qquad \boldsymbol{S}(\boldsymbol{Q}, \boldsymbol{p}) \tag{16}$$

represents a canonical transformation, i.e., one that preserves the commutation rules. Namely, as will be soon shown, (16) can be put into the form:

$$\boldsymbol{Q}_k = \boldsymbol{T}^{-1} \boldsymbol{q}_K \boldsymbol{T}, \qquad \boldsymbol{P}_k = \boldsymbol{T}^{-1} \boldsymbol{p}_K \boldsymbol{T}, \qquad [\boldsymbol{T}(\boldsymbol{Q}, \boldsymbol{p})], \tag{16′}$$

when $\boldsymbol{T}$ and $\boldsymbol{S}$ can be chosen in a certain corresponding way. According to **Jordan** (^{1}), the connection between the function $\boldsymbol{S}$ and the function $\boldsymbol{T}$ that mediates the same transformation is as follows:

Let $\boldsymbol{S}(\boldsymbol{Q}, \boldsymbol{q})$ be developed in the form:

(^{1}) **P. Jordan**, Zeit. Phys. **38** (1926), pp. 513.

$$\boldsymbol{S}\,(\boldsymbol{Q}, \boldsymbol{q}) = \sum_{s=1}^{m} \boldsymbol{f}_s\,(\boldsymbol{Q}) \cdot \boldsymbol{g}_s(\boldsymbol{p}) \tag{17}$$

as the sum of m products, the *first* factor of which is a function that depends upon only $\boldsymbol{Q}$ and the *second* of which depends upon only $\boldsymbol{p}$. The factors do not commute with each other. With **Pauli**, one now defines the exponential function $e^{\boldsymbol{x}\,\boldsymbol{y}}$, in which the factors $\boldsymbol{x}$ and $\boldsymbol{y}$ are functions of quantum quantities that do not commute, by the series:

$$e^{\boldsymbol{x}\,\boldsymbol{y}} = \sum_{r=0}^{\infty} \frac{\boldsymbol{x}^r \boldsymbol{y}^r}{r!}. \tag{18}$$

Its generalization is the exponential function:

$$\exp\left[\sum_{s=1}^{m} \boldsymbol{x}_s\, \boldsymbol{y}_s\right] = \sum_{r_1=1}^{\infty}\sum_{r_2=0}^{\infty}\cdots\sum_{r_m=0}^{\infty} \frac{\boldsymbol{x}_1^{r_1}\cdots \boldsymbol{x}_m^{r_m}\cdot \boldsymbol{y}_1^{r_1}\cdots \boldsymbol{y}_m^{r_m}}{r_1!\,r_2!\cdots r_m!}, \tag{18′}$$

in which the $\boldsymbol{x}$-factors shall precede the $\boldsymbol{y}$-factors. It is only when the $\boldsymbol{x}$ commute with the $\boldsymbol{y}$ that one can also write:

$$e^{\sum \boldsymbol{x}_s\,\boldsymbol{y}_s} = e^{\boldsymbol{x}_1\boldsymbol{y}_1}\cdot e^{\boldsymbol{x}_2\boldsymbol{y}_2}\cdots e^{\boldsymbol{x}_m\boldsymbol{y}_m}.$$

Like **Jordan**, we would now like to show that the transformed $\boldsymbol{T}\,(\boldsymbol{q}, \boldsymbol{p})$, which mediates the same transformation as $\boldsymbol{S}\,(\boldsymbol{Q}, \boldsymbol{p})$, is given by the exponential function:

$$\boldsymbol{T}\,(\boldsymbol{q}, \boldsymbol{p}) = \exp\left[\frac{2i\pi}{h}\boldsymbol{S}_*(\boldsymbol{Q}, \boldsymbol{p})\right], \tag{19}$$

in which $\boldsymbol{S}_*(\boldsymbol{Q}, \boldsymbol{p})$ shall mean [cf., (17)]:

$$\boldsymbol{S}_*(\boldsymbol{Q}, \boldsymbol{p}) = \boldsymbol{S}\,(\boldsymbol{Q}, \boldsymbol{p}) - \sum_K \boldsymbol{Q}_K\ \boldsymbol{p}_K = \sum_{s=1}^{m} \boldsymbol{f}_s(\boldsymbol{Q})\,\boldsymbol{g}_s(\boldsymbol{p}) - \sum_K \boldsymbol{Q}_K\ \boldsymbol{p}_K\,. \tag{20}$$

In order to prove the formula (19), one observes that it will follow from the general differential formulas (4):

$$\boldsymbol{T}\,\boldsymbol{q}_K - \boldsymbol{q}_K\,\boldsymbol{T} = \frac{h}{2i\pi}\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{p}_K}, \qquad \boldsymbol{p}_K\,\boldsymbol{T} - \boldsymbol{T}\,\boldsymbol{p}_K = \frac{h}{2i\pi}\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{q}_K}$$

that:

$$\boldsymbol{P}_K = \boldsymbol{T}^{-1}\boldsymbol{p}_K\boldsymbol{T} = \boldsymbol{p}_K + \frac{h}{2i\pi}\,\boldsymbol{T}^{-1}\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{q}_K}, \qquad \boldsymbol{Q}_K = \boldsymbol{T}^{-1}\boldsymbol{q}_K\boldsymbol{T} = \boldsymbol{q}_K - \frac{h}{2i\pi}\,\boldsymbol{T}^{-1}\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{p}_K}.$$

If one replaces the differential quotients of $\boldsymbol{T}$ in the right-hand sides of these with the functions that follow from (19) and observes that according to (8′), one will have:

$$\boldsymbol{T}^{-1}\boldsymbol{f}_s(\boldsymbol{q})\boldsymbol{T} = \boldsymbol{f}_s(\boldsymbol{Q}), \qquad \boldsymbol{T}^{-1}\frac{\partial \boldsymbol{f}_s(\boldsymbol{q})}{\partial \boldsymbol{q}_K}\boldsymbol{T} = \frac{\partial \boldsymbol{f}_s(\boldsymbol{Q})}{\partial \boldsymbol{Q}_K},$$

then one will, in fact, get the transformation formulas (16) precisely.

The special case of the *point transformation* is characterized by the generator:

$$\boldsymbol{S}(\boldsymbol{Q},\boldsymbol{p}) - \sum_K \boldsymbol{v}_K(\boldsymbol{Q})\,\boldsymbol{p}_K\,, \tag{21}$$

so

$$\boldsymbol{q}_K - \frac{\partial \boldsymbol{S}}{\partial \boldsymbol{p}_K} = \boldsymbol{v}_K(\boldsymbol{Q}), \qquad \boldsymbol{P}_K = \frac{\partial \boldsymbol{S}}{\partial \boldsymbol{Q}_K} = \sum_K \frac{\partial \boldsymbol{v}_K(\boldsymbol{Q})}{\partial \boldsymbol{Q}_K}\cdot \boldsymbol{p}_K\,.$$

Here, one has:

$$\boldsymbol{S}_*(\boldsymbol{Q},\boldsymbol{p}) = \sum_K [\boldsymbol{v}_K(\boldsymbol{Q}) - \boldsymbol{Q}_K]\cdot \boldsymbol{p}_K\,,$$

so:

$$\boldsymbol{T}(\boldsymbol{q},\boldsymbol{p}) = \exp\left[\frac{2i\pi}{h}\boldsymbol{S}_*(\boldsymbol{q},\boldsymbol{p})\right] = \exp\left[\frac{2i\pi}{h}\sum_K [\boldsymbol{v}_K(\boldsymbol{q}) - \boldsymbol{q}_K]\cdot \boldsymbol{p}_K\right]. \tag{21′}$$

The transformed T, which is mediated by the point transformation $\boldsymbol{q}_K = \boldsymbol{v}_K(\boldsymbol{Q})$, is then a more complicated transcendental function of the $\boldsymbol{q}$ and $\boldsymbol{p}$.

62. Transformations in wave mechanics. – In wave mechanics, one deals with the solution of the differential equation:

$$\left\{H\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - E, \psi(q)\right\} = 0\,, \tag{22}$$

which can be made easier in some situations by the introduction of new coordinates. It is the *point* transformations $q = q(Q)$ that first come under consideration, by which the fundamental equation will go to the form:

$$\left\{H\left(q(Q), \frac{h}{2i\pi}\frac{dQ}{dq}\frac{\partial}{\partial q}\right) - E, \psi(q(Q))\right\} = 0\,.$$

In its place, we would like to write more concisely:

$$\left\{H_*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right) - E, \psi^*(Q)\right\} = 0\,, \tag{23}$$

with the eigenfunctions:

$$\psi_n^*(Q) = \psi_n(q(Q)) \, ,$$

and the same eigenfunctions E_n as in the original equation (22).

General canonical transformations in wave mechanics were first treated by **F. London** (1) following the example of **Jordan**'s transformation in quantum mechanics. Starting from an arbitrary function $T(q, p)$, one defines the operator $T\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)$ and sets:

$$Q_K = T^{-1} q_K T \, , \qquad \frac{h}{2i\pi}\frac{\partial}{\partial Q_K} = T^{-1}\frac{h}{2i\pi}\frac{\partial}{\partial q_K} T \, . \tag{24}$$

In that way, an operator $F^*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)$ will go to the operator [cf., (8′)]:

$$F^*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right) = T^{-1} F_*\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) T = F\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) . \tag{24′}$$

The **Schrödinger** equation for the desired eigenfunctions $\psi^*(Q)$ and eigenvalues E^* of the **Hamiltonian** operator then reads:

$$\left\{H_*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right) - E^*, \psi^*(Q)\right\} = 0 \, . \tag{25}$$

We will now show that this possesses the same eigenvalues as the original equation (22) (which was just now shown for only point transformations). To that end, we use (24′) to write:

$$\left\{T^{-1} H_*\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) T - E, \psi(q)\right\} = 0 \, ,$$

instead of (22) and apply the operator T:

$$\left\{H_*\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) T - T\,E, \psi(q)\right\} = 0 \, .$$

If we set:

$$T \cdot \psi(q) = \psi^*(q) \, , \qquad \text{i.e.,} \qquad \psi(q) = T^{-1}\psi^*(q) \, , \tag{26}$$

in that then we will get:

---

(1) **F. London**, Zeit. Phys. **40** (1926), pp. 193.

$$\left\{H_*\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - E, \psi^*(q)\right\} = 0\ ,$$

and finally when one replaces the symbols q with Q :

$$\left\{H_*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right) - E, \psi^*(Q)\right\} = 0\ . \tag{27}$$

If one compares (27) with (25) then that will show the agreement between the *eigenvalue* E and E^*. The connection between the *eigenfunctions* that is mediated by the transforming T is given by (26):

$$\begin{aligned} \psi_n^*(Q) &= T\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)\cdot\psi_n(Q), \\ \psi_n(q) &= T^{-1}\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)\cdot\psi_n^*(q) \end{aligned} \qquad \text{where } Q_K = T^{-1}q_K T\ ,\ P_K = T^{-1}p_K T\ . \tag{28}$$

The situation in which the domain of existence for the q does not need to coincide with that of the Q requires special consideration. Let q be, say, a length coordinates $-\infty \leq q \leq +\infty$, and let Q be an angle $0 \leq Q \leq 2\pi$. How that difficulty in this special case will resolve itself automatically will be shown in the example in no. **64**.

When the transforming T includes the argument p in a fractional or negative power, the transformations will give rise to non-whole-number and negative "differentiations" $\frac{h}{2i\pi}\frac{\partial^n}{\partial q^n}$. We shall not go into their meaning, but only remark that the formula:

$$\frac{d^n e^{kx}}{dx^n} = k^n\, e^{kx}$$

that occurs in the transformation to angle variables should also be adapted to fractional and negative n.

(28) gives and answer to the following question: Suppose that we know the eigenfunction $\psi_n(q)$, which we would like to refer to as the *probability amplitude*, that belongs to the energy value E_n , and that for the energy value E_n , the coordinates q_K are found in certain locations q_K. How large is the probability amplitude $\psi_n^*(Q)$ that the functions $Q_K(q, p)$ will have certain values Q_K ? The answer (28) assumes the one has found the transforming T that will mediate the canonical transformation $Q\,P \to q\,p$.

We now consider the special case in which the coordinate Q_1, which appears as an argument in the probability amplitude $\psi_n^*(Q_1, Q_2, \ldots, Q_N)$, is equal to the **Hamiltonian** function $Q_1 = H\,(q,$

p). One will then have $Q_1 = T^{-1} q_1 T = H(q, p)$, so $q_1 T = T H$, and with the use of $E_n \psi_n(q) = H \psi_n(q)$, (28) will imply that:

$$E_n \psi_n^*(q) = E_n T \psi_n(q) = T E_n \psi_n(q) = T H \psi_n(q) = q_1 T \psi_n(Q) .$$

The equation $E_n \psi_n^*(q) = E_n \cdot T \psi_n = q_1 \cdot T \psi_n$, The probability amplitude $\psi_n^*(q)$ will have a non-vanishing value only for $q_1 = E_n$, and $\psi_n^*(Q)$ can have a non-vanishing functional value only for $Q_1 = E_n$. In other words, $\psi_n^*(Q)$ will have *an infinitely-steep maximum* there and vanish everywhere else.

The connection between the operator $T\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)$ that mediates a canonical transformation:

$$Q_K = T^{-1} q_K T , \qquad \frac{h}{2i\pi}\frac{\partial}{\partial Q_K} = T^{-1} \frac{h}{2i\pi}\frac{\partial}{\partial q_K} T \tag{29}$$

and the operator that generates the same transformation with the equations of transformation:

$$q_K = \frac{\partial S\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)}{\partial p_K} , \qquad \frac{\partial S\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)}{\partial Q_K} = \frac{h}{2i\pi}\frac{\partial}{\partial Q_K} \tag{30}$$

is significant. If the connection between the functions $\boldsymbol{T}(\boldsymbol{q}, \boldsymbol{p})$ and $\boldsymbol{S}(\boldsymbol{Q}, \boldsymbol{p})$ were given according to **Jordan** then according to **London**, one would get the corresponding connection between the operators T and S from:

$$T\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) = \exp\left[\frac{2i\pi}{h} S^*\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)\right], \tag{31}$$

in which:

$$S\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) = \sum_{s=1}^{m} f_s(Q) \cdot g_s\left(\frac{h}{2i\pi}\frac{\partial}{\partial q}\right) \quad \text{and} \quad S^* = S - \sum_K Q_K \cdot \frac{h}{2i\pi}\frac{\partial}{\partial q_K} , \tag{32}$$

as in (17), (20). In the special case of the *point transformation* $q_K = v_K(Q)$, (21′) will correspondingly become:

$$T\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) = \exp\left\{\sum_K [v_K(q) - q_K] \cdot \frac{\partial}{\partial q_K}\right\}.$$

In this case, when one uses (26) and (18′), one will have:

$$\psi^*(q) = T\,\psi(q) = \exp\left\{\sum_K [v_K(q) - q_K]\cdot\frac{\partial}{\partial q_K}\right\}\psi(q)\,.$$

However, that is a **Taylor** series:

$$\psi^*(q) = \psi(q_1 + [v_1(q) - q_1],\, q_2 + [v_2(q) - q_2],\, \ldots) = \psi(v(q))\,.$$

If one now writes the symbol Q instead of q in the relation thus-found $\psi^*(q) = \psi(v(q))$, so:

$$\psi_n^*(Q) = \psi_n(v(Q)) \qquad \text{(point transformation),} \tag{33}$$

then one will have the result that for *point transformations* $q_K = v_K(Q)$, the eigenfunction $\psi_n^*(Q)$ in the new variables will go to the eigenfunction $\psi(q)$ in the old variables by the *substitution* $q_K = v_K(Q)$ as a special case of the general connection (26):

$$\psi_n^*(Q) = T\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)\psi_n(Q) \qquad \text{(general canonical transformation).} \tag{34}$$

63. Angle variables in wave mechanics. – Let us now succeed in introducing new coordinates w such that the original **Hamiltonian** operator $H\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)$ will become an operator $H^*\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right)$ that does not include w as an argument, but only $\partial / \partial w$, and which can be developed in powers of $\partial / \partial w$ with coefficients that are independent of w, i.e., *constants:*

$$H^*\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right)\psi^*(w) = \sum_\tau A_\tau\left(\frac{h}{2i\pi}\right)^\tau \frac{\partial^\tau}{\partial w^\tau}\psi^*(w) = E\,\psi^*(w)\,.$$

(For the sake of simplicity, we shall now consider a system with *only one* degree of freedom.) That eigenvalue problem can then be solved immediately by the Ansatz $\psi^*(w) = Be^{\beta w}$, and has the eigenvalue:

$$E(\beta) = \sum_\tau A_\tau\left(\frac{h}{2i\pi}\right)^\tau\cdot\beta^\tau = H^*\left(\frac{h\beta}{2i\pi}\right), \tag{35}$$

in which B is determined from the boundary conditions. In classical mechanics, the *angle variables* are the conjugate coordinates w and impulses J for which the **Hamiltonian** function will be a function of only the J, while the position coordinates are periodic in the w with a period of 2π. In wave mechanics, corresponding angle variables are distinguished in such a way that the

Hamiltonian operator does not depend upon w, but only on $\partial / \partial w$ (to arbitrarily high order), while the eigenfunctions shall be periodic in w as functions of the position coordinates with period 2π. Now, the latter demand fixes the boundary condition that was left open in (35) and ultimately gives:

$$\psi_n^*(w) = \frac{1}{\sqrt{2\pi}} e^{inw}, \quad E_n = H^*\left(\frac{hn}{2\pi}\right) \qquad (n = 0, 1, 2, \ldots) \tag{36}$$

as the solution of the eigenvalue problem. In that way, the factor $B = 1/\sqrt{2\pi}$ is determined by the normalization:

$$\int_0^{2\pi} \psi_n^*(w) \cdot \tilde{\psi}_n^*(w)\, dw = 1 .$$

Now, one often demands that in addition to $\psi_n^*(w) = \frac{1}{\sqrt{2\pi}} e^{inw}$, the eigenfunctions $\psi_n(q)$ should belong to the same eigenvalue E_n when they are given in the original coordinates. According to **London**, that problem can be solved as follows: Let $S(q, J)$ be the generator of the transformation:

$$w = \frac{\partial S\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right)}{\partial J}, \qquad \frac{h}{2i\pi}\frac{\partial}{\partial q} = \frac{\partial S\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right)}{\partial q}$$

that mediates the transition from q to w, and let it be represented is the form of a series:

$$S\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right) = \sum_{s=1}^{m} f_s(q) \cdot g_s\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right), \qquad S^* = \sum_{s=0}^{m} f_s(q) \cdot g_s\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right),$$

with $f_0(q) = -q$, $g_0\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right) = \frac{h}{2i\pi}\frac{\partial}{\partial w}$. If one knows the transforming $T_*\left(w, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right)$ that belongs to the same transformation then from (26) and (18′), the desired representation will be given by:

$$\psi_n(w) = T_*\left(w, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right) \cdot e^{inw} = \exp\left[\frac{2i\pi}{h}\sum_{s=0}^{m} f_s(w) \cdot g_s\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right)\right] e^{inw}$$

$$= \sum_{r_1=0}^{\infty} \cdots \sum_{r_m=0}^{\infty} \frac{\prod_{\nu=0}^{m}\left(\frac{2i\pi}{h}\right)^{r_\nu} f_\nu^{r_\nu}(w) \cdot \prod_{\nu=0}^{m}\left(\frac{2i\pi}{h}\right)^{r_\nu} g_\nu^{r_\nu}(w)}{\prod_{\nu=0}^{\infty} r_\nu !} e^{inw} .$$

However, when the first operation $g_\nu^{r_\nu}\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right)$ that results is applied to e^{inw}, it will be identical to the *multiplication* by $g_\nu^{r_\nu}\left(n\frac{h}{2\pi}\right)\cdot e^{inw}$, and one will then get the total result:

$$\psi_n(w) = T_*\left(w, \frac{nh}{2\pi}\right)\cdot e^{inw} = \exp\left[\frac{2i\pi}{h}\sum_{s=0}^{m} f_s(w)\cdot g_s\left(\frac{nh}{2\pi}\right)\right]e^{inw}$$

$$= \exp\left[-inw + \frac{2i\pi}{h}\sum_{s=0}^{m} f_s(w)\cdot g_s\left(\frac{nh}{2\pi}\right)\right]e^{inw} = \exp\left[\frac{2i\pi}{h}S\left(w, \frac{nh}{2\pi}\right)\right],$$

or finally when one formally replaces the symbol w with q :

$$\psi_n(q) = \exp\left[\frac{2i\pi}{h}S\left(q, \frac{nh}{2\pi}\right)\right]. \tag{37}$$

Since the quantities q and $n\,h$ *commute* in S (all of the differential operations have already been performed), an ordinary exponential function will be present in (37), such that one no longer needs to appeal to the series representation of S at all. That is why **London**'s (^{1}) result (37) is more meaningful in the eigenvalue problems because it shows how to represent the eigenfunctions $\psi_n(q)$ of a quantum-mechanical system (here, it will initially have *one* degree of freedom) in terms of arbitrary coordinates q as images of the exponential (trigonometric, resp.) functions, which are the prototype for all oscillatory processes.

Moreover, **London**'s formula $\psi_n = e^{2i\pi S/h}$ represents the exact limit of an approximation procedure for solving the wave equation that was given before by **Wentzel** and **Brillouin**. **Wentzel** (^{2}) employed the **de Broglie** Ansatz in order to solve the wave equation:

$$\psi(q) = e^{\frac{2i\pi}{h}\int y\,dq}, \qquad \text{so} \qquad y = \frac{h}{2i\pi}\frac{\psi'}{\psi}, \tag{38}$$

which will reduce the wave equation to a **Riccati** differential equation in y (q). Its solution can then be replaced with a series in increasing powers of h :

$$y(q) = \sum_{s=0}^{\infty}\left(\frac{h}{2i\pi}\right)^s\cdot y_s(q)\ ,$$

(^{1}) **F. London**, Zeit. Phys. **40** (1926), pp. 193.

(^{2}) **G. Wentzel**, Zeit. Phys. **38** (1926), pp. 518.

whose zeroth approximation y_0 (q) corresponds to classical mechanics $y_0 = p$, $\int y_0\,dq = \int p\,dq$, while the higher approximation terms give a successive approximation to quantum mechanics. If one then calculates $\oint y\,dq$ around a closed path in a region where all n zeroes (i.e., nodes) of the eigenfunction ψ_n lie, so in which $y = \frac{h}{2i\pi}\frac{\psi'}{\psi}$ possesses n poles with residue h, then that will the values of the integral as:

$$\oint y\,dq = n\,h \quad (n = \text{whole number} = \text{number of nodes}).$$

According to **Wentzel**, one will then get the **Sommerfeld-Wilson** quantum condition $\oint p\,dq = nh$ as the zeroth approximation to the fundamental equation of wave mechanics for the determination of the eigenvalues. Moreover, as **Wentzel** shows in several examples, the series development will truncate after a few terms in many cases, such that one will get a finite number of approximations to the rigorous solution of the eigenvalue problem.

Guided by the optical-mechanical analogy, **Brillouin** (^{1}) considered y (q) to be a wave function of a generalized *eikonal* S by the Ansatz:

$$y\,(q) = e^{2i\pi S(q)/h}, \tag{39}$$

and developed S (q) in powers of h. In that way, the classical action function appeared as a first term (the factor of h^0). The connection between the *eikonal* and the *action function* that generates the the canonical transformation to angle variables was clarified in **London**'s examination [cf., (37).

64. Harmonic oscillator. – **London**'s (^{2}) example of the harmonic oscillator will serve to illustrate the canonical transformations in wave mechanics.

The wave equation of the oscillator:

$$-\frac{h^2}{8\pi^2}\frac{\partial^2\psi(q)}{\partial q^2} + \tfrac{1}{2}\nu^2 q^2\,\psi(q) = E\,\psi(q)\,, \tag{40}$$

will be canonically transformed by means of the generator:

$$S_1\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial x}\right) = i\left(\frac{\nu}{2}q^2 + \frac{h}{2i\pi}\sqrt{2\nu}\,q\frac{\partial}{\partial x} - \frac{h^2}{8\pi^2}\frac{\partial^2}{\partial x^2}\right)$$

into a variable x:

---

(1) **L. Brillouin**, C. R. Acad. Sci. Paris **183** (July 1926), pp. 24.

(2) **F. London**, Zeit. Phys. **40** (1926), pp. 193.

$$\frac{h\nu}{2\pi} x \frac{\partial \psi^*(x)}{\partial x} + \frac{h\nu}{4\pi} \psi^*(x) = E\psi^*(x),$$

and then transform that to the angle variable *w* by means of:

$$S_2\left(x, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right) = -\frac{h}{2\pi} \cdot \ln x \cdot \frac{\partial}{\partial w} \qquad \text{(point transformation)}$$

so one will have:

$$\frac{h\nu}{2i\pi} x \frac{\partial \psi^{**}(w)}{\partial w} + \frac{h\nu}{4\pi} \psi^{**}(w) = E\psi^{**}(w).$$

One easily verifies that $\psi_n^{**}(w) = e^{inw}$ $\left[\psi_n^*(x) = \exp\left[\frac{2i\pi}{h} S_2\left(x, \frac{nh}{2\pi}\right)\right] = x^n\right]$, with the eigenvalues $E_n = \left(n+\tfrac{1}{2}\right)\frac{h\nu}{2\pi}$, are single-valued solutions to the last two equations. It is the transition to the eigenfunctions $\psi_n(q)$ that is of interest here. One finds from (31), (32) with the help of the generators S_1 by the action of the operator:

$$\begin{aligned} T\left(x, \frac{h}{2i\pi}\frac{\partial}{\partial x}\right) &= \exp\left[\frac{2i\pi}{h} S_1\left(x, \frac{h}{2i\pi}\frac{\partial}{\partial x}\right) - x\frac{\partial}{\partial x}\right] \\ &= \exp\left[-\frac{2\nu\pi}{2h}x^2 + i\sqrt{2\nu}\, x\frac{\partial}{\partial x} + \frac{h}{4\pi}\frac{\partial^2}{\partial x^2} - x\frac{\partial}{\partial x}\right] \\ &= \exp\left(-\frac{2\nu\pi}{2h}x^2\right) \cdot \exp\left(x(i\sqrt{2\nu}-1)\frac{\partial}{\partial x}\right) \cdot \exp\left(\frac{h}{4\pi}\frac{\partial^2}{\partial x^2}\right) \end{aligned}$$

on $\psi_n^*(x) = x^n$. One then gets $\psi_n(x)$ step-wise:

$$\begin{aligned} \psi_n(x) &= \left[T\left(x, \frac{h}{2i\pi}\frac{\partial}{\partial x}\right), x^n\right] \\ &= \exp\left(-\frac{2\nu\pi}{2h}x^2\right) \cdot \exp\left(x(i\sqrt{2\nu}-1)\frac{\partial}{\partial x}\right) \\ &\cdot \left(\frac{h}{4\pi}\right)^{n/2} \left\{\left(\sqrt{\frac{4\pi}{h}}\,x\right)^n + \frac{n(n-1)}{1!}\left(\sqrt{\frac{4\pi}{h}}\,x\right)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!}\left(\sqrt{\frac{4\pi}{h}}\,x\right)^{n-4} + \cdots\right\} \end{aligned}$$

The operation $\exp\left(x(i\sqrt{2\nu}-1)\frac{\partial}{\partial x} \right)$ that must now be performed is the substitution $x \to i\sqrt{2\nu}\,x$:

$$\psi_n(x) = \exp^{-\frac{2\nu\pi}{2h}x^2} \cdot \left(\frac{h}{4\pi} \right)^{n/2}$$
$$\cdot i^n \left\{ \left(2\sqrt{\frac{2\pi\nu}{h}}\,x \right)^n - \frac{n(n-1)}{1!}\left(2\sqrt{\frac{2\pi\nu}{h}}\,x \right)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!}\left(2\sqrt{\frac{2\pi\nu}{h}}\,x \right)^{n-4} - + \cdots \right\}.$$

In that way, one will have found a new system of orthogonal functions, even if they are not the ones that are known already as the ***Hermite*** *polynomials:*

$$\psi_n(x) = \text{const.} \cdot \exp^{-\frac{2\nu\pi}{2h}x^2} \cdot H_n\left(\sqrt{\frac{2\pi\nu}{h}}\,x \right).$$

65. Matrix algebra. – If $H(p, q)$ is a certain symmetrized Hamiltonian function, while E_n and $\psi_n(q)$ are the eigenvalues and normalized eigenfunctions of the associated **Schrödinger** equation $\{H - E, \psi\} = 0$, and if $F(q, p)$ is an arbitrary function, moreover, the one understands the matrix components of the function $F(q, p)$ relative to the **Hamiltonian** function $H(q, p)$ to mean the quantities:

$$\mathfrak{F}_{mn} = \int \tilde{\psi}_m(q)\, F\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q} \right) \psi_n(q)\, dv. \tag{41}$$

The schema:

$$\mathfrak{F} = \left\{ \begin{matrix} \mathfrak{F}_{11} & \mathfrak{F}_{12} & \mathfrak{F}_{13} & \cdots \\ \mathfrak{F}_{21} & \mathfrak{F}_{22} & \mathfrak{F}_{23} & \cdots \\ \mathfrak{F}_{31} & \mathfrak{F}_{32} & \mathfrak{F}_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix} \right\}$$

is the matrix $\mathfrak{F}$ with the components $\mathfrak{F}_{mn}$. The definition (41) is identical to the following one: Let $\boldsymbol{F}(\boldsymbol{q}, \boldsymbol{p})$ be the *transformer* from the $\boldsymbol{q}$, $\boldsymbol{p}$ to other canonical variables $\boldsymbol{Q}$, $\boldsymbol{P}$, namely:

$$\boldsymbol{Q}_K = \boldsymbol{F}^{-1}\boldsymbol{q}_K\,\boldsymbol{F}, \qquad \boldsymbol{P}_K = \boldsymbol{F}^{-1}\boldsymbol{p}_K\,\boldsymbol{F}. \tag{42}$$

$H(q, p)$ will then transform into $H_*(Q, P)$ with the same eigenvalues E_n as before, but the transformed eigenfunctions $\psi^*(Q)$. According to (28), the connection between the $\psi_n(q)$ and the $\psi_n^*(Q)$ is then:

$$\left.\begin{aligned} \psi_n(q) &= F^{-1}\left(q, \frac{h}{2i\pi} q\right)\psi_n^*(q), \quad \text{i.e.,} \quad \psi_n = \{F^{-1}, \psi_n^*\}, \\ \psi_n^*(Q) &= F\left(q, \frac{h}{2i\pi} q\right)\psi_n(Q), \quad \text{i.e.,} \quad \psi_n^* = \{F, \psi_n\}. \end{aligned}\right\} \tag{43}$$

If one now makes the development Ansatz:

$$\psi_n^*(q) = \sum_m \mathfrak{F}_{mn}\, \psi_m(q)$$

then one will calculate the coefficients $\mathfrak{F}_{mn}$ in the development to be:

$$\mathfrak{F}_{mn} = \int \tilde{\psi}_m(q)\, \psi_n^*(q)\, dv = \int \tilde{\psi}_m(q)\, F\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)\psi_n(q)\, dv,$$

which are equal to the quantities that were defined in (41). That is, *the matrix elements* $\mathfrak{F}_{mn}$ *of an arbitrary function F* (*q*, *p*) *are the coefficients in the development of the series:*

$$\psi_n^*(q) = \sum_m \mathfrak{F}_{mn}\, \psi_m(q). \tag{44}$$

We shall now derive the basic rules of calculation for the metric components (1). From the definition (41), one will the *addition law* for two different functions $F(q, p)$ and $G(q, p)$:

$$(\mathfrak{F} \pm \mathfrak{G})_{mn} = \mathfrak{F}_{mn} \pm \mathfrak{G}_{mn}, \qquad \text{and furthermore} \qquad (\mathfrak{F} + \mathfrak{G})_{mn} = (\mathfrak{G} + \mathfrak{F})_{mn}. \tag{45}$$

In order to derive the law of multiplication, i.e., to express the components $(\mathfrak{F}\mathfrak{G})_{mn}$ in terms of the components F_{kl} and G_{kl}, we use *F*, *G*, and *FG* as the transformer for the following transformation (42) to new variables and new eigenfunctions:

$$\begin{aligned} Q'_K &= F^{-1} q_K F, & P'_K &= F^{-1} p_K F, & &\text{with} \quad \psi'_n(Q'), \\ Q''_K &= G^{-1} q_K G, & P''_K &= G^{-1} p_K G, & &\text{with} \quad \psi''_n(Q''), \\ Q'''_K &= (FG)^{-1} q_K (FG), & P'''_K &= (FG)^{-1} p_K (FG), & &\text{with} \quad \psi'''_n(Q'''). \end{aligned}$$

With the help of (43), (44), one will then have:

---

(1) In connection with this, cf., **F. London**, Zeit. Phys. **40** (1926), pp. 193.

$$\begin{aligned} \psi'_l &= \{F,\ \psi_l\} &&= \sum_m \mathfrak{F}_{ml}\,\psi_m\,, \\ \psi''_l &= \{G,\ \psi_l\} &&= \sum_l \mathfrak{G}_{l\,n}\,\psi_l\,, \\ \psi'''_n &= \{F\,G,\ \psi_n\} &&= \sum_m (\mathfrak{F}\,\mathfrak{G})_{mn}\,\psi_m\,. \end{aligned}$$

With the use of the first two of those three equations, we can convert the last of them into:

$$\sum_m (\mathfrak{F}\,\mathfrak{G})_{mn}\,\psi_m = \{F\,G,\ \psi_n\} = \{F, \sum_l \mathfrak{G}_{l\,n}\,\psi_l\} = \sum_l \mathfrak{G}_{l\,n}\,\{F, \psi_l\} = \sum_l \sum_m \mathfrak{G}_{l\,n}\,\mathfrak{F}_{ml}\,\psi_m$$

which implies the *multiplication rule:*

$$(\mathfrak{F}\ \mathfrak{G})_{mn} = \sum_l \mathfrak{F}_{ml}\,\mathfrak{G}_{l\,n}\ . \tag{46}$$

It has the form of the known rule by which the elements of the product determinant $\mathfrak{F} \cdot \mathfrak{G}$ can be composed from the elements of the determinants $\mathfrak{F}$ and $\mathfrak{G}$. In that way, $\mathfrak{F} \cdot \mathfrak{G}$ is not always equal to $\mathfrak{G} \cdot \mathfrak{F}$, but from (45) one does have $\mathfrak{F} + \mathfrak{G} = \mathfrak{G} + \mathfrak{F}$. It follows further from (46) that:

$$(\mathfrak{F}\ \mathfrak{G}\ \mathfrak{H})_{mn} = \sum_l \sum_k \mathfrak{F}_{ml}\,\mathfrak{G}_{l\,n}\,\mathfrak{H}_{kn}\ , \qquad \text{etc.} \tag{46}$$

Now that the components of the matrices that result from arbitrary functions $F(q, p)$ and $G(q, p)$ by addition and multiplication have been expressed in terms of the components of $\mathfrak{F}$ and $\mathfrak{G}$ themselves, the question still remains of what the general non-vanishing values of the components of $(\mathfrak{F}\mathfrak{G} - \mathfrak{G}\mathfrak{F})_{mn}$ would be. As a special case, we consider $\mathfrak{F} = \mathfrak{p}_K$, $\mathfrak{G} = \mathfrak{q}_L$, and with the use of (41), we will get:

$$(\mathfrak{p}_K\ \mathfrak{q}_L - \mathfrak{q}_L\ \mathfrak{p}_K)_{mn} = \int \tilde{\psi}_m(q) \left(\frac{h}{2i\,\pi} \frac{\partial}{\partial q_K}\, q_L - q_L\, \frac{h}{2i\,\pi} \frac{\partial}{\partial q_K} \right) \psi_n(q) \cdot dv\ = 0 \qquad \text{for} \qquad K \neq L,$$

$$= \frac{h}{2i\,\pi} \int \tilde{\psi}_m(q)\,\psi_n(q) \cdot dv\ = \begin{cases} \dfrac{h}{2i\pi} & \text{for} \quad m = n \\ 0 & \text{for} \quad m \neq n \end{cases} \qquad \text{for} \qquad K = L\ .$$

If one understands δ_{ik} to have the value 0 for $i \neq k$ and the value 1 for $i = k$ then the latter result can be written:

$$(\mathfrak{p}_K\ \mathfrak{q}_L - \mathfrak{q}_L\ \mathfrak{p}_K)_{mn} = 0 \qquad \text{for} \qquad K \neq L,$$

$$(\mathfrak{p}_K\ \mathfrak{q}_L - \mathfrak{q}_L\ \mathfrak{p}_K)_{mn} = \frac{h}{2i\,\pi}\,\delta_{mn}\,.$$

All of the rules of association, distribution, and commutation can be summarized as follows:

$$\left.\begin{array}{c}(\mathfrak{F}+\mathfrak{G})+\mathfrak{H}=\mathfrak{F}+(\mathfrak{G}+\mathfrak{H}), \quad (\mathfrak{F}\cdot\mathfrak{G})\cdot\mathfrak{H}=\mathfrak{F}\cdot(\mathfrak{G}\cdot\mathfrak{H}), \\ \mathfrak{F}(\mathfrak{G}+\mathfrak{H})=\mathfrak{F}\cdot\mathfrak{G}+\mathfrak{F}\cdot\mathfrak{H}, \\ \mathfrak{F}=0 \quad \text{or} \quad \mathfrak{G}=0 \quad \text{when} \quad \mathfrak{F}\cdot\mathfrak{G}=0, \\ (\mathfrak{F}+\mathfrak{G})-(\mathfrak{G}+\mathfrak{F})=0, \quad \text{but} \quad (\mathfrak{F}\cdot\mathfrak{G})-(\mathfrak{G}\cdot\mathfrak{F}) \quad \text{is ordinarily} \neq 0,\end{array}\right\} \tag{47}$$

and indeed, one will have, in particular:

$$\left.\begin{array}{ll}\mathfrak{p}_K\,\mathfrak{p}_L-\mathfrak{p}_L\,\mathfrak{p}_K=0, & \mathfrak{q}_K\,\mathfrak{q}_L-\mathfrak{q}_L\,\mathfrak{q}_K=0, \\ \mathfrak{p}_K\,\mathfrak{q}_L-\mathfrak{q}_L\,\mathfrak{p}_K=0 \quad \text{for } K\neq L, & \mathfrak{p}_K\,\mathfrak{q}_K-\mathfrak{q}_K\,\mathfrak{p}_K=\dfrac{h}{2i\pi}\mathbf{1},\end{array}\right\} \tag{48}$$

in which **1** is understood to mean the matrix with the components δ_{mn} (“identity matrix”). (47) are the usual laws of determinant algebra, and (48) gives the *commutation rules* for the special determinants $\mathfrak{p}_K$ and $\mathfrak{q}_K$. (47), (48) are formally the same rules that were true in no. **44** for the operators $\boldsymbol{F}$ ($\boldsymbol{q}$, $\boldsymbol{p}$), …, and in particular for the $\boldsymbol{q}_K$ and $\boldsymbol{p}_K$ themselves. *Operator calculations* (calculating with “quantum quantities”) and *matrix calculations* (calculating with determinants) are formally identical when one observes the elementary commutation rules, although one must observe that the *components* of the matrices are always defined only with respect to a certain **Hamiltonian** $H(p, q)$ [with respect to any two of its eigenfunctions $\psi_m(q)$ and $\psi_n(q)$, resp.].

Since the **Schrödinger** equation says that $H(q, p)\ \psi_n(q) = E_n\ \psi_n(q)$, the matrix element of the **Hamiltonian** function will be:

$$\mathfrak{H}_{mn} = \int \tilde{\psi}_m\, H\left(q_1, \frac{h}{2i\,\pi}\frac{\partial}{\partial q}\right)\psi_n\, dv = \int \tilde{\psi}_m\, E_n\, \psi_n\, dv = E_n \cdot \delta_{mn}\,. \tag{49}$$

$\mathfrak{H}_{mn}$ is a a *diagonal matrix,* i.e., only the elements in the diagonal $m = n$ can be non-zero. In general, due to (41), one will have:

$$\mathfrak{F}_{mn} = \tilde{\mathfrak{F}}_{nm}\,, \tag{49′}$$

moreover.

We now ask how we might express the matrix components of the function $\dot{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{p})$ in terms of those of $\boldsymbol{x}(\boldsymbol{q}, \boldsymbol{p})$ itself. Now, in (5′), we defined:

$$\dot{\boldsymbol{x}}(\boldsymbol{q}, \boldsymbol{p}) = \frac{2i\,\pi}{h}(\boldsymbol{H}\,\boldsymbol{x} - \boldsymbol{x}\,\boldsymbol{H})\,,$$

so due to (46) and (49), we will then get:

$$\dot{\mathfrak{x}}_{mn} = \frac{2i\pi}{h}(\mathfrak{H}\;\mathfrak{x} - \mathfrak{x}\;\mathfrak{H})_{mn} = \frac{2i\pi}{h}\sum_k (\mathfrak{H}_{mk}\,\mathfrak{x}_{kn} - \mathfrak{x}_{mk}\,\mathfrak{H}_{kn})$$

when we set:

$$\left.\begin{aligned} &= \frac{2i\pi}{h}(\mathfrak{H}_{mm} - \mathfrak{H}_{nn})\,\mathfrak{x}_{mn} = 2i\pi\,\nu^{(mn)}\cdot\mathfrak{x}_{mn}, \\ &\frac{\mathfrak{H}_{mm} - \mathfrak{H}_{nn}}{h} = \frac{E_m - E_n}{h} = \nu^{(mn)}. \end{aligned}\right\} \tag{50}$$

Division can be introduced when we define the matrix F^{-1} by the equation:

$$F \cdot F^{-1} = F^{-1}F = 1, \qquad \text{i.e.,} \qquad \sum_k \mathfrak{F}_{mk}\,\mathfrak{F}^{-1}_{kn} = \sum_k \mathfrak{F}^{-1}_{mk}\,\mathfrak{F}_{kn} = \delta_{mn}\,. \tag{51$'$}$$

One can regard $\psi\,(q)$ as a vector field with infinitely-many “components” $\psi_1\,(q)$, $\psi_1\,(q)$, $\psi_3\,(q)$, … along the “axes” 1, 2, 3, … The $\psi^*_n(q)$ can then be regarded as the components of that vector field $\psi\,(q)$ along new *rotated* axes 1^*, 2^*, 3^*, …, because the “coordinate transformation: of the vector components:

$$\psi^*_n(q) = \sum_m \mathfrak{F}_{mn}\,\psi_m(q)$$

will be mediated by a coefficient schema $\mathfrak{F}_{mn}$, which fulfills the *orthogonality relations:*

$$\sum_l \tilde{\mathfrak{F}}_{lj}\,\mathfrak{F}_{lk} = \delta_{jk} \quad (= 1 \text{ for } j = k, = 0 \text{ for } j \neq k), \tag{51$''$}$$

which one proves by the following sequence of equations:

$$\delta_{jk} = \int \tilde{\psi}^*_j\,\psi^*_k\,dv = \int \sum_l \tilde{\mathfrak{F}}_{lj}\,\tilde{\psi}_l \cdot \sum_i \mathfrak{F}_{ik}\psi_i\,dv = \sum_l\sum_i \tilde{\mathfrak{F}}_{lj}\,\mathfrak{F}_{ik}\,\delta_{li} = \sum_l \tilde{\mathfrak{F}}_{lj}\,\mathfrak{F}_{lk}\,.$$

66. Invariance of the matrix components. – It is essential for the possibility of physically interpreting the matrix components $\mathfrak{G}_{mn}$ of a function $G\,(q, p)$ relative to a certain Hamiltonian function $H\,(p, q)$ that the values $\mathfrak{G}_{mn}$ must not change when one goes from the q, p to the new canonical variables Q, P, and in that way go from $H\,(q, p)$ to $H_*(Q, P)$. **F. London** (^{1}) proved that as follows: Define:

(^{1}) **F. London**, Zeit. Phys. **40** (1926), pp. 193.

$$\mathfrak{G}_{mn} = \int \tilde{\psi}_m(q)\, G\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q} \right) \psi_n(q)\, dq\,,$$

$$\mathfrak{G}^*_{mn} = \int \tilde{\psi}^*_m(q)\, G_*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q} \right) \psi^*_n(Q)\, dQ\,,$$

under which $H(p, q)$ and $\psi_n(q)$ will become $H^*(Q, P)$ and $\psi^*_n(Q)$, resp., under the transformation (42), and $G(q, p) = G^*(Q, P)$, so $G_*(q, p) = F\,G(q, p)\,F^{-1}$. Now, with the help of (43), (44), here we have:

$$\begin{aligned} \mathfrak{G}^*_{mn} &= \int \tilde{\psi}^*_m(q) \cdot G_*\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q} \right) \cdot \psi^*_n(q)\, dq\,, \\ &= \int \sum_k \tilde{\mathfrak{F}}_{im}\, \tilde{\psi}_i(q) \cdot FGF^{-1} \cdot \sum_j \mathfrak{F}_{jn}\, \psi_j(q)\, dq \\ &= \sum_i \sum_j \tilde{\mathfrak{F}}_{im}\, (\mathfrak{F}\,\mathfrak{G}\,\mathfrak{F}^{-1})_{ij} \cdot \mathfrak{F}_{jn} = \sum_i \sum_j \sum_k \sum_l \tilde{\mathfrak{F}}_{im}\, \mathfrak{F}_{ik}\, \mathfrak{G}_{kl}\, \mathfrak{F}^{-1}_{lj}\, \mathfrak{F}_{jn}\,, \end{aligned}$$

and finally, from (51′), (51″), one will have:

$$\mathfrak{G}^*_{mn} = \sum_k \sum_l \delta_{mk}\, \mathfrak{G}_{kl}\, \delta_{ln} = \mathfrak{G}_{mn}\,,$$

which proves the invariance, i.e., the matrix components of $G(q, p)$ relative to the $\psi(q)$ are equal to the matrix components of $G^*(Q, P)$ relative to $\psi^*(Q)$. At the same time, the eigenvalues E_n are also equal to E^*_n, as was shown before in no. **62**.

67. Matrix mechanics. – The solution to the quantum-mechanical problem in operator calculus (no. **60**) consists of finding a canonical transformation of the variables $\boldsymbol{q}$, $\boldsymbol{p}$ to new variables $\boldsymbol{w}$, $\boldsymbol{J}$ for a given function $\boldsymbol{H}(\boldsymbol{q}, \boldsymbol{p})$ such that:

(*a*) The Hamiltonian function in the new variables depends upon only J.

(*b*) The variables $\boldsymbol{q}$ and $\boldsymbol{p}$ can be represented as series:

$$\boldsymbol{q}_K = \sum_\tau \boldsymbol{q}^{(\tau)}_K(\boldsymbol{J}) \cdot e^{i(\tau w)}\,, \qquad \boldsymbol{p}_K = \sum_\tau \boldsymbol{p}^{(\tau)}_K(\boldsymbol{J}) \cdot e^{i(\tau w)}\,.$$

The values E_n of the energy will then be equal to $H^*\left(\frac{nh}{2\pi}\right)$. However, one will also arrive at the same energy values with the use of only the *matrix components:* For a given function $H(q, p)$, when one uses the rule of calculation:

$$(\mathfrak{x}\ \mathfrak{y})_{mn} = \sum_k \mathfrak{x}_{mk}\ \mathfrak{y}_{kn}\ , \tag{52}$$

one seeks a system of values:

$$(\mathfrak{p}_K)_{mn}\ , \qquad (\mathfrak{q}_K)_{mn}$$

that fulfills the canonical conditions:

$$\left.\begin{aligned} &(\mathfrak{p}_K\,\mathfrak{q}_L - \mathfrak{q}_L\,\mathfrak{p}_K)_{mn} = 0 \quad \text{for} \quad K \neq L\,, \\ &(\mathfrak{p}_K\,\mathfrak{q}_K - \mathfrak{q}_K\,\mathfrak{p}_K)_{mn} = \frac{h}{2i\pi}\delta_{mn}\,, \\ &(\mathfrak{p}_K\,\mathfrak{p}_L - \mathfrak{p}_L\,\mathfrak{p}_K)_{mn} = 0, \quad (\mathfrak{q}_K\,\mathfrak{q}_L - \mathfrak{q}_L\,\mathfrak{q}_K)_{mn} = 0\,, \end{aligned}\right\} \tag{53}$$

which further:

(*a*) makes the Hamiltonian matrix a diagonal matrix:

$$\mathfrak{H}_{mn} = 0 \quad \text{for} \quad m \neq n\,, \tag{54}$$

and

(b) satisfies the condition (49′):

$$\mathfrak{q}_{mn} = \tilde{\mathfrak{q}}_{nm}\,, \qquad \mathfrak{p}_{mn} = \tilde{\mathfrak{p}}_{nm}\,. \tag{55}$$

One will arrive at such a system of values when one initially starts from any system of quantities $(\mathfrak{q}'_K)_{mn}$, $(\mathfrak{p}'_K)_{mn}$ that at least satisfies the rules (52), (53), (55), and then goes to variables q_K, p_K that also satisfy the conditions (54) by a suitable canonical transformation:

$$(q'_K)_{mn} = (T^{-1}q_K T)_{mn}\,, \qquad (p'_K)_{mn} = (T^{-1}p_K T)_{mn}\,.$$

Moreover, according to equation (50), which can also write as a matrix relation:

$$\dot{\mathfrak{x}} = \frac{2i\pi}{h}(\mathfrak{H}\ \mathfrak{x} - \mathfrak{x}\ \mathfrak{H})\,, \tag{56}$$

and in particular, for $\mathfrak{x} = \mathfrak{q}_K$ or $\mathfrak{x} = \mathfrak{p}_K$:

$$\dot{\mathfrak{q}}_K = \frac{2i\pi}{h}(\mathfrak{H}\ \mathfrak{q}_K - \mathfrak{q}_K\ \mathfrak{H})\,, \quad \text{or} \quad \dot{\mathfrak{p}}_K = \frac{2i\pi}{h}(\mathfrak{H}\ \mathfrak{p}_K - \mathfrak{p}_K\ \mathfrak{H}), \quad \text{resp.} \tag{57}$$

Those identities are the *canonical equations of motion* for matrix mechanics, just as (6′) were the ones for operator mechanics. For $\mathfrak{x} = \mathfrak{H}$, one will get:

$$\dot{\mathfrak{H}} = 0\,, \qquad \text{i.e.,} \qquad (\dot{\mathfrak{H}})_{mn} = 0\,. \tag{58}$$

That is the *law of conservation of energy* in matrix mechanics.

The possibility that one can solve the quantum-mechanical problem of determining the energy eigenvalues E_m for a given system $H(q, p)$ with the use of only *matrix components* and the relations between them was regarded by **Heisenberg**, **Born**, and **Jordan** (^{1}), who were the founders of matrix mechanics, as an epistemological advantage over **Schrödinger**'s theory, which addressed the determination of eigenfunctions ψ_m of partial differential equations. That is because the forms of the eigenfunctions depend upon the random choice of coordinate system q, p or Q, P or w, J, etc., so the eigenfunctions cannot be represented as *physically*-invariant quantities, as opposed to the invariant matrix components.

One interprets the components H_{mm} as energy values, while the components x_{mn} are the amplitude functions $x(q, p)$ as the amplitudes that are definitive of the radiation of frequency $\nu^{(mn)} = (E_m - E_n) / h$,

68. Generalized probability amplitudes. – In **Schrödinger**'s wave mechanics, the eigenfunction $\psi_k(q)$ was interpreted as the probability *amplitude*, while $|\ \psi_k(q)\ |^2\ dv$ was interpreted as the *probability* that the system occupied the coordinate interval dv for a given energy value E_k . Now, with **Jordan** (^{2}), one can ask, more generally, what the probability amplitude $\psi(q, \beta)$ might be for the system to occupy the position q for given values β_K of any *parameters* that are introduced as functions $\beta_K(q, p)$. [Up to now, we have ordinarily taken β_1 to be the energy function $\beta_1(q, p) = -H(q, p)$.] Once more, let:

$$q_1, \ldots, q_N, \qquad p_1, \ldots, p_N, \qquad \text{or briefly} \qquad q_K, p_K$$

be the coordinates and impulses (i.e., operators), which satisfy the canonical commutation rules:

$$p_K\, q_K - q_K\, p_K = \frac{h}{2i\pi}\,, \qquad p_K\, q_L - q_L\, p_K = 0 \qquad \text{for} \qquad K \neq L\,. \tag{59}$$

Furthermore, let:

$$\beta_K = \beta_K(q, p)\,, \qquad \alpha_K = \alpha_K(q, p) \tag{60}$$

be new transformed coordinates and canonically-conjugate impulses, and indeed let the kinematic connection between the p, q and the α, β be given by the transforming $T(p, q)$ in the form:

$$\beta_K = \beta_K(q, p) = T^{-1} q_K T\,, \qquad \alpha_K = \alpha_K(q, p) = T^{-1} p_K T \qquad (\alpha, \beta \to p, q). \tag{61}$$

(^{1}) **W. Heisenberg**, Zeit. Phys. **33** (1925), pp. 879; **M. Born** and **P. Jordan**, *ibidem*, **34** (1925), pp. 856; **Heisenberg**, **Born**, and **Jordan**, *ibidem*, **35** (1926), pp. 557.

(^{2}) **P. Jordan**, "Über eine neue Begründung der Quantenmechanik," Göttinger Nachr. (1926), pp. 161.

One now asks what a generalized probability amplitude $\varphi(q, \beta)$ might be that solves the following system of $2N$ differential equations:

$$\begin{aligned} &\left\{\alpha_K\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + \frac{h}{2i\pi}\frac{\partial}{\partial \beta_K}, \varphi(q,\beta)\right\} = 0, \\ &\left\{\beta_K\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - \beta_K, \qquad \varphi(q,\beta)\right\} = 0 \end{aligned} \qquad (K = 1, 2, \ldots, N). \qquad (62)$$

As will be soon shown, those equations for $\varphi(q, \beta)$ *first of all*, include the **Schrödinger** equation as a special case, and *secondly*, the function $\varphi(q, \beta)$ possesses properties that characterize it as a probability amplitude and $|\varphi|^2 \, dv$ as a probability that the q_K will be found in the interval dv for a given value of β. The fact that here the *one* function $\varphi(q, \beta)$ can be simultaneously subject to $2N$ differential equations is possible only because the α_K and β_K define a canonical system and emerge from the q, p by means of the *one* function $T(q, p)$.

The equations do, in fact, first include the **Schrödinger** equation as a special case. Namely, if one takes the new coordinate $\beta_1(p, q)$ to be minus the Hamiltonian function $-H(p, q)$, and accordingly takes $\beta_1 = -E$, $\alpha_1 = +t$, then the pair of equations for $K = 1$ will go to:

$$\left.\begin{aligned} &\left\{t\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - \frac{h}{2i\pi}\frac{\partial}{\partial E}, \varphi(q_1, q_2, \ldots, E, \beta_2, \ldots)\right\} = 0, \\ &\left\{-H\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + E, \qquad \varphi(q_1, q_2, \ldots, E, \beta_2, \ldots)\right\} = 0. \end{aligned}\right\} \qquad (63)$$

The latter is **Schrödinger**'s equation of oscillation, which is the analogue of the classical energy equation $H(p, q) - E = 0$. With the Ansatz $\varphi = e^{-2i\pi S/h}$, the former will go to an equation for S:

$$\left\{t\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + \frac{\partial S(q_1, q_2, \ldots, E, \beta_2, \ldots)}{\partial E}, \quad \varphi(q_1, q_2, \ldots, E, \beta_2, \ldots)\right\} = 0, \qquad (64)$$

which is an analogue of the classical equation $t + \partial S / \partial E = 0$ for the action function S. However, along with the pair of equations (63), $N - 1$ further pairs of equations will appear for $K = 2, 3, \ldots, N$ that are the classical counterparts to the **Jacobi** equations $\alpha_K - \partial S / \partial \beta_K = 0$ [no. **12**, (42)].

In order to exhibit $\varphi(q, \beta)$ as secondly a type of probability amplitude, **Jordan** proved the following *theorem:* If $\varphi(q, \beta)$ is the solution to equations (62) that belongs to the transformation (61) $\alpha, \beta \rightarrow p, q$ then $\varphi(Q, q)$ will further be the solution to the equations:

$$\left.\begin{aligned}&\left\{p_K\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)+\frac{h}{2i\pi}\frac{\partial}{\partial q_K},\psi(Q,q)\right\}=0,\\&\left\{p_K\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)-q_K,\qquad\psi(Q,q)\right\}=0\end{aligned}\right\}\tag{65}$$

that belongs to a transformation:

$$q_K = q_K(Q,P) = S^{-1}Q_K S,\quad p_K = p_K(Q,P) = S^{-1}P_K S\qquad (p, q \to P, Q),\tag{66}$$

and ultimately $\Phi(Q, \beta)$ is the solution to the equations:

$$\left.\begin{aligned}&\left\{A_K\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)+\frac{h}{2i\pi}\frac{\partial}{\partial \beta_K},\Phi(Q,\beta)\right\}=0,\\&\left\{B_K\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)-\beta_K,\qquad\Phi(Q,\beta)\right\}=0\end{aligned}\right\}\tag{67}$$

that belongs to the direct transformation ($\alpha, \beta \to P, Q$):

$$\left.\begin{aligned}\alpha_K &= \alpha_K(q,p)=\alpha_K(S^{-1}P\,S,S^{-1}Q\,S)=S^{-1}\alpha_K(Q,P)\,S=A_K(P,Q),\\\beta_K &= \beta_K(q,p)=\beta_K(S^{-1}P\,S,S^{-1}Q\,S)=S^{-1}\beta_K(Q,P)\,S=B_K(P,Q)\end{aligned}\right\}\tag{67′}$$

then, as can be proved immediately:

$$dQ\cdot\Phi(Q,\beta) = dQ\int\psi(Q,q)\cdot\varphi(q,\beta)\,dq\,.\tag{68}$$

However, that is the law for combining the *probability amplitude:*

$$\Phi(Q,\beta)\qquad\text{from}\qquad\psi(Q,\beta)\text{ and }\varphi(Q,\beta)\,.$$

Since it is the *amplitudes*, and not the probabilities themselves, that are being combined here according to the rule (68), **Jordan** saw in (68) an *interference* of probabilities.

In order to prove (68), one uses the operator $S\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)$ in order to put (67) into the form $\{S^{-1}S\cdot[\cdots]\cdot S^{-1}S,\Phi\}=0$, i.e., from (67′), in the form:

$$\left.\begin{aligned}&\left\{S^{-1}\left[\alpha_K\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)+\frac{h}{2i\pi}\frac{\partial}{\partial\beta_K}\right]S,\Phi(Q,\beta)\right\}=0,\\&\left\{S^{-1}\left[\beta_K\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)-\beta_K\right]S,\qquad\Phi(Q,\beta)\right\}=0.\end{aligned}\right\}\tag{68′}$$

If one now switches the symbol q with Q in (62) then a comparison with (68′) will show that $S\,\Phi(Q,\beta) = \varphi(Q,\beta)$, so one will have:

$$\Phi(Q,\beta) = S^{-1}\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)\varphi(Q,\beta)\ ,\tag{69}$$

and due to the generality of the parameter β, that equation is somewhat more general than equation (28) of **London**'s transformation theory, which takes the energy parameter E_n to be a particular value for β. One can now multiply (69) by an arbitrary function $f(\beta)$ on the right and left and integrate over β:

$$\int d\beta\,\Phi(Q,\beta)\,f(\beta) = S^{-1}\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)\int d\beta\,\varphi(Q,\beta)\,f(\beta),$$

such that it will ultimately result that:

$$S\left(Q,\frac{h}{2i\pi}\frac{\partial}{\partial Q}\right)\int d\beta\,\Phi(Q,\beta)\,f(\beta) = \int d\beta\,\varphi(Q,\beta)\,f(\beta)\ .\tag{69′}$$

One now chooses $S(Q, P)$ to be equal to $T^{-1}(Q,P)$, in particular, i.e., one transforms $\alpha, \beta \to p, q$ and back again $p, q \to \alpha, \beta$. One will especially have:

$$B_K(P, Q) = T(Q, P)\,\beta_K(Q, P)\,T(Q, P) = Q_K\,, \qquad \text{and likewise} \quad A_K(Q, P) = PK\,,$$

in place of (67) and with the use of (64), such that (67) will imply that (we write χ instead of Φ, in particular):

$$\left.\begin{aligned}&\left\{\frac{h}{2i\pi}\frac{\partial}{\partial Q_K}+\frac{h}{2i\pi}\frac{\partial}{\partial Q_K},\chi(Q,\beta)\right\}=0,\\&\{\qquad Q_K-\beta_K,\qquad \chi(Q,\beta)\}=0.\end{aligned}\right\}\tag{70}$$

Due to the first of those equations, $\chi(Q, \beta)$ will depend upon only the differences $Q_K - \beta_K$:

$$\chi(Q, \beta) = \Phi'(Q_1 - \beta_1, \ldots, Q_N - \beta_N)\ .$$

Due to the second one, Φ' can be non-zero only when $Q_1 = \beta_1$, and at the same time $Q_2 = \beta_2$, etc. In that way, χ can be normalized such that $\int d\beta\, \chi(Q,\beta)$ has the value 1. (Only the location $\beta_1 = Q_1$, $\beta_2 = Q_2$, … contributes to the integral.). If one now substitutes T^{-1} for S in equation (69′), and accordingly substitutes χ for Φ, then when one writes q instead of Q and appends an arbitrary function f, it will become:

$$T^{-1}\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)\cdot f(q) = \int d\beta \cdot \varphi(q,\beta)\, f(\beta)\,, \tag{71}$$

which is a relation between the function T that transforms α, β to p, q and the probability amplitude φ (q, β) that is associated with it by way of (62). The latter appears in (71) as a "generating function" of the operator $T^{-1}\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right)$. Therefore, (71) is a consequence of (61), (62). When one appeals to an arbitrary function F, it will correspondingly follow from (65), (66). If one sets the arbitrary function here $F(Q) = \varphi(q, \beta)$ then one will get the following formula from (69):

$$\Phi\,(Q, \beta) = \int dq \cdot \psi\,(Q, q) \cdot \varphi(q, \beta)\ ,$$

which proves (68).

69. Angle variables. – The search for the solution $\varphi(q, \beta)$ of (62) will then be complete when one succeeds in going over to *angle variables*: One might find a transforming S (Q, P) such that (67′) will assume the special form:

$$A_K(Q, P) = -\,Q_K\,, \qquad B_K(Q, P) = P_K\,. \tag{73}$$

We then write $Q_K = w_K$, $P_K = J_K$ (angle variables). In that way, (67) will become:

$$\left.\begin{aligned} &\left\{-\,w_K + \frac{h}{2i\pi}\frac{\partial}{\partial \beta_K}, \Phi\,(w,\beta)\right\} = 0,\\ &\left\{\frac{h}{2i\pi}\frac{\partial}{\partial \beta_K} - \beta_K, \quad \Phi\,(w,\beta)\right\} = 0. \end{aligned}\right\} \tag{73′}$$

Those equations will now be solved immediately by:

$$\Phi\,(w, \beta) = \exp\left[\frac{2i\pi}{h}\sum_N w_K\,\beta_K\right], \tag{74}$$

and from (69), one will get the solution to the original equations (62):

$$\varphi(w, \beta) = S\left(w, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right)\Phi(w, \beta)\,, \tag{74′}$$

in which the symbol w can be replaced with q. If $S\left(w, \frac{h}{2i\pi}\frac{\partial}{\partial w}\right)$ can be represented as a series:

$$S = \sum_m u_m(w)\cdot v_m\left(\frac{h}{2i\pi}\frac{\partial}{\partial w}\right) \tag{75}$$

then one will get the solution from (74′) (cf., no. **63**) in the explicit form:

$$\varphi(q, \beta) = \sum_m u_m(q)\cdot v_m(\beta)\cdot \exp\left[\frac{2i\pi}{h}\sum_N q_K\,\beta_K\right]. \tag{75′}$$

70. Sharpness in the observation of physical quantities. – The answer to a very general question is included in equations (62) for the probability amplitude $\varphi(q, \beta)$:

$$\begin{aligned} &\left\{\alpha_K\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) + \frac{h}{2i\pi}\frac{\partial}{\partial \beta_K}, \varphi(q,\beta)\right\} = 0\,, \\ &\left\{\beta_K\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - \beta_K, \qquad \varphi(q,\beta)\right\} = 0. \end{aligned} \tag{76}$$

The **Schrödinger equation:**

$$\left\{H\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - H, \psi(q)\right\} = 0 \tag{77}$$

would be solved for the eigenvalues H_n of the parameter H by the eigenfunctions $\psi_n(q)$ that belong to $H(q, p)$, relative to which the matrix of $H(p, q)$ will become a *diagonal matrix* ($\mathfrak{H}_{mn} = 0$ for $m \neq n$, $\mathfrak{H}_{nn} = H_n$). In an entirely corresponding way, one will have solved:

$$\left\{\beta\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) - \beta, \varphi(q)\right\} = 0 \tag{78}$$

for the eigenvalues β_1, β_2, … by way of eigenfunctions $\varphi_1(q)$, $\varphi_2(q)$, …, relative to which the function $\beta(q, p)$ will give a *diagonal matrix:*

$$\beta_{mn} = \int \tilde{\varphi}_m \, \beta\left(q, \frac{h}{2i\pi}\frac{\partial}{\partial q}\right) \varphi_n \cdot dv = \begin{cases} 0 & \text{for} \quad m \neq n, \\ \beta_n & \text{for} \quad m = n. \end{cases} \tag{79}$$

Just as the eigenfunctions $\psi_n(q)$ of (77) give the probability amplitudes for a coordinate value q to be occupied for a given $H = H_n$, so do the eigenfunctions $\varphi_n(q)$ of (78) give the probability amplitudes that a coordinate value q will be occupied for a given value of $\beta(q, p) = \beta_n$.

The probability amplitude $\varphi_n^*(Q)$ for the functions $Q_K(q, p)$ to have the values Q_K for given β_n is then determined from the differential equation that belongs to the transforming function $\beta^*(Q, P) = \beta(q, p)$:

$$\left\{\beta^*\left(Q, \frac{h}{2i\pi}\frac{\partial}{\partial Q}\right) - \beta_n, \varphi_n^*(Q)\right\} = 0\,. \tag{80}$$

If one takes $Q_1(q, p)$ to be the function $\beta(q, p)$ itself, in particular, then as was shown in the corresponding case in the middle of no. **62**, $|\varphi_n^*(Q)|^2$ will have a non-vanishing value in the form of an *infinitely-sharp maximum* only for $Q_1 = \beta_n$. By contrast, if all $Q_K(q, p)$ are different from $\beta(q, p)$ then $|\varphi_n^*(Q)|^2$ will possess maxima that are more or less sharp for certain values of the Q_K. Finally, if $Q_1 = \alpha(q, p)$ and $P_1 = \beta(q, p)$, in particular, in which α means the canonical conjugate to β, then it will follow from the transformation $\beta^*(Q, P) = \beta(q, p)$, i.e., $\beta^*(\alpha, Q_2, \ldots, \beta, P_2, \ldots) = \beta$, that the operator β^* is equal to $\frac{h}{2i\pi}\frac{\partial}{\partial \alpha}$, such that from (80), one will have:

$$\left\{\frac{h}{2i\pi}\frac{\partial}{\partial \alpha} - \beta_n, \varphi_n^*(\alpha, Q_2, \ldots)\right\} = 0\,. \tag{81}$$

That equation will be solved by the probability amplitude:

$$\varphi_n^*(\alpha, Q_2, \ldots) = \omega_n(Q_2, \ldots) \cdot e^{2i\pi\alpha\beta_n/h}, \tag{82}$$

with the probability function:

$$|\varphi_n^*(Q)|^2 = |\,\omega_n(Q_2, \ldots)\,|^2, \tag{83}$$

which no longer includes the argument α at all, but is equally-large for all values of α. We have then found that:

For a given value β_n of the function $\beta(q, p)$, the probability that $\beta(q, p)$ itself has a certain value β will be concentrated around the sharp value $\beta = \beta_n$. By contrast, the probability that the quantity $\alpha(q, p)$ that is conjugate to β will have a certain value α is distributed uniformly over the entire range of variability for α with no preference given to any special value of α. On the other hand, any other function $Q_K(q, p)$ will possess a probability function with *maxima* that are more or less *indistinct* for any given value $\beta(q, p) = \beta_n$. Physically, that means that for a given $\beta(q, p) = \beta_n$, the determination of the quantity $Q_K(q, p)$ will be possible with only greater or lesser uncertainty, which is characterized by the course of the function $|\varphi_n^*(Q)|^2$. On the contrary, a determination of the quantity $\alpha(q, p)$ that is conjugate to β is not possible at all.

In that way, the question of finding the ultimate limits of the precision that are obtainable in the observation of any physical quantity $Q_K(q, p)$ for a fixed value β_n of another physical quantity $\beta(q, p)$ will become a quantum-theoretic problem.