QUANTUM MECHANICS AND ASYMPTOTIC SERIES†

BY G. D. BIRKHOFF

PART I.

1. *Introduction.* In its bold primary outline the program of quantum mechanics in the Schrödinger form runs as follows:

(A) Set up the Hamiltonian equations of the atomic system (nucleus+electrons) on a classical basis:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial y_i}; \quad \frac{dy_i}{dt} = -\frac{\partial H}{\partial x_i}, \qquad (i = 1, \cdots, n),$$

where x_i, y_i are ordinary rectangular coordinates and momenta, and $H(x_1, \cdots, x_n; y_1, \cdots, y_n)$ is the total energy. The associated Hamilton-Jacobi partial differential equation is then

$$\frac{\partial S}{\partial t} + H\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right) = 0.$$

(B) Write down the corresponding homogeneous linear partial differential equation (the Schrödinger wave equation):

$$\frac{1}{\lambda}\frac{\partial \psi}{\partial t} + H\left(x_1, \cdots, x_n; \frac{1}{\lambda}\frac{\partial}{\partial x_1}, \cdots, \frac{1}{\lambda}\frac{\partial}{\partial x_n}\right)\psi = 0,$$

where the operational symbols in H appear on the right hand side of the individual terms of H, and where $\lambda = 2\pi i/h$, if h is Planck's constant.

(C) Write

$$\psi = e^{-\lambda E t}\psi^*(x_1, \cdots, x_n),$$

thus obtaining the linear differential equation (written in operational form)

$$(H - E)\psi^* = 0,$$

and determine the characteristic solutions, $\psi_1^*, \psi_2^*, \cdots$, for which the ψ_i^* vanish at infinity in such wise that

† An address delivered at Chicago, June 20, 1933, before the Society and Section A of the American Association for the Advancement of Science. In connection with the first part of this lecture, see two notes in the Proceedings of the National Academy of Sciences for March and April, 1933.

$$\int |\psi_i^*|^2 dx_1 \cdots dx_n$$

is finite. The corresponding $E_1, E_2, \cdots$ then prescribe the possible "energy levels," so that the possible "spectral frequencies," ν_{mn}, are those given by the formula

$$h\nu_{mn} = E_m - E_n, \qquad (E_m > E_n),$$

in accordance with the Planck-Einstein law.

Thus the program begins by relating the physical problem to a special linear boundary value problem of classical type. In its further development the aim is to obtain a complete account of atomic properties which is in accord with this starting point. While there has been extraordinary progress, there can be little doubt that complete success of the program is hardly to be hoped for.

My purpose today is to lay before you a tentative answer to two important mathematical questions raised by the primary program itself.

Firstly, what is the *mathematical* significance of the Schrödinger wave equation in its relation to the Hamiltonian equations? My answer will be given in terms of classic formal processes connected with asymptotic series. It is true that the theoretical physicist has obtained a kind of "deduction" of the wave equation on the basis of the analogy between the wave theory of light and the elementary optical theory (wave and particle theory). But I hope to bring out more clearly the true inwardness of the Schrödinger wave equation as a purely mathematical entity.†

Secondly, the form of the wave equation which arises in practice is usually reducible by means of separation of variables to an ordinary differential equation essentially of the following type:

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}(E - V(x))\psi = 0,$$

where the function $V(x)$ is defined in a certain interval in which $E - V(x)$ changes sign. For the determination of the character-

† For references to the important earlier work of de Broglie, L. Brillouin, Schrödinger, and Dirac, see my notes cited above.

istic numbers and functions, certain asymptotic series have been employed by Wentzel, Brillouin and Kramers.† It is, however, an open question as to whether or not their methods are justifiable. However, in an important recent paper bearing directly upon the points at issue, Langer‡ announces that he intends later to give a general discussion of this and similar questions. My intention here is to outline a simple justification of the Wentzel-Brillouin-Kramers method on the basis of a slight extension of some earlier results of my own.

Thus asymptotic series play a central role in what I have to say today. I would not be surprised if such series were found ultimately to be of importance in other aspects of quantum mechanics; for example, in connection with the proper formulation of Heisenberg's uncertainty principle.

2. *Linear Equations and Asymptotic Series.* Let

$$L(\psi, \lambda) = 0 \tag{1}$$

be any linear homogeneous differential equation in the dependent variable ψ and the independent variables $x_1, x_2, \cdots, x_n$, and involving λ, where λ is a large parameter. This equation will be ordinary or partial according as $n=1$ or $n>1$. Let us suppose that the coefficients of ψ and of its derivatives in $L(\psi, \lambda)$ are analytic in $x_1, \cdots, x_n$, and λ, and expansible in convergent power series in $1/\lambda$ for $|\lambda|>\Lambda$.

Now under these circumstances it has frequently been found that the differentiation of certain solutions ψ, as well as of their various derivatives with respect to x_i, is asymptotically equivalent to multiplication by $\lambda\partial S/\partial x_i$, where S is a suitable function of $x_1, \cdots, x_n$. For instance in the simple case $n=1$ of an equation

$$L(\psi, \lambda) \equiv \frac{1}{\lambda^2}\frac{d^2\psi}{dx^2} + \psi = 0 \tag{1a}$$

with solutions $e^{\pm\lambda ix}$, we have $d\psi/dx = \pm\lambda i\psi$, so that $S = \pm ix$ in

† For the principal references see a note by J. L. Dunham, *On Wentzel-Brillouin-Kramers' method of solving the wave equation*, Physical Review, vol. 41 (Sept. 15, 1932).

‡ R. E. Langer, *On the asymptotic solutions of differential equations, with an application to the Bessel functions of large complex order*, Transactions of this Society, vol. 34 (1932), pp. 447–480.

this case. Again, in the case of the Fourier's equation with independent variable taken as λx,

$$(1b)\qquad L(\psi, \lambda) \equiv \frac{1}{\lambda^2}\frac{d^2\psi}{dx^2} + \frac{1}{\lambda^2}\frac{1}{x}\frac{d\psi}{dx} + \psi = 0,$$

we have $S = \pm ix$ as before for suitable solutions.

In consequence of such an asymptotic relationship we are led to write

$$\psi \sim e^{\lambda S(x_1, \cdots, x_n)} v_0(x_1, \cdots, x_n)$$

as a first approximation to ψ, and thence successively to an asymptotic series for ψ:

$$\psi \sim e^{\lambda S}\left(v_0 + \frac{v_1}{\lambda} + \frac{v_2}{\lambda^2} + \cdots\right),$$

where $v_0, v_1, \cdots$ are functions of $x_1, \cdots, x_n$. The precise test for such an asymptotic series solution is of course that when it is substituted for ψ in $L(\psi, \lambda)$, with the indicated differentiations carried out and the coefficients of like powers of λ collected according to the usual formal rules, the expression $L(\psi, \lambda)$ reduces identically to 0. It is not to be expected in general that such a series converges and yields an actual solution, although this may occur in special cases. Obviously it may be assumed that v_0 does not vanish identically, since otherwise we could remove a factor $1/\lambda$ from the solution.

We shall term S a "phase function," and any corresponding v_0 an "amplitude function" of S. The reasons for this designation will appear subsequently.

We propose first to outline some fundamental facts concerning this classical formal process which have apparently escaped attention. For this purpose we find it convenient to introduce the modified differential operators

$$\frac{\partial^{[1]}F}{\partial x_i} = \frac{1}{\lambda}\frac{\partial F}{\partial x_i}, \quad \frac{\partial^{[2]}F}{\partial x_i \partial x_j} = \frac{1}{\lambda^2}\frac{\partial^2 F}{\partial x_i \partial x_j}, \text{ etc.}\dagger$$

In the two special cases noted above this notation allows us to write the equations (1a), (1b) as follows:

† See my paper, Transactions of this Society, vol. 9 (1908), pp. 219–231.

$$\frac{d^{[2]}\psi}{dx^2} + \psi = 0; \qquad \frac{d^{[2]}\psi}{dx^2} + \frac{1}{\lambda}\frac{1}{x}\frac{d^{[1]}\psi}{dx} + \psi = 0,$$

while, more generally, $L(\psi, \lambda)$ may be written as a power series in $1/\lambda$,

$$L(\psi, \lambda) \equiv L_0(\psi) + \frac{1}{\lambda}L_1(\psi) + \cdots, \tag{2}$$

where $L_i(\psi)$, $(i=0, 1, \cdots)$, are linear homogeneous expressions in ψ, $\partial^{[1]}\psi/\partial x_i$, $\partial^{[2]}\psi/\partial x_i\partial x_j$, etc., and where we may assume that $L_0(\psi) \not\equiv 0$, since we may always multiply through by a suitable power of λ. Evidently, then, we may write

$$L_i(\psi) = \xi^{(i)}\psi + \sum_j \xi_j^{(i)} \frac{\partial^{[1]}\psi}{\partial x_j} + \sum_{jk} \xi_{jk}^{(i)} \frac{\partial^{[2]}\psi}{\partial x_j \partial x_k} + \cdots. \tag{3}$$

Furthermore, on account of the interchangeability of the order of differentiation, we may assume

$$\xi_{jk}^{(i)} = \xi_{jk}^{(i)}; \quad \xi_{jkl}^{(i)} = \xi_{jlk}^{(i)} = \xi_{klj}^{(i)} = \cdots; \cdots. \tag{4}$$

The order of the expression $L(\psi, \lambda)$ is evidently that of the highest order of L_0, $L_1, \cdots$, say m. We shall assume that m is the actual order of $L_0(\psi)$, and shall term $L_0(\psi)$ the "principal part" of $L(\psi, \lambda)$; in dealing with L_0 we shall omit the superscripts in referring to $\xi^{(0)}$, $\xi_i^{(0)}, \cdots$.

Later on we shall have something to say concerning the existence of actual solutions of (1) corresponding to such asymptotic series solutions in the special case $m=2$. For the present we shall make only the following heuristic remarks. (1) Solutions asymptotically represented by such series solutions will exist for suitably restricted ranges of the variables $x_1, \cdots, x_n$. (2) These actual solutions are not uniquely determined by the series solutions; for example, if ψ is so represented, so also will $(1+e^{-\lambda})\psi$ be if λ is real and positive. (3) The solution $c_1\psi_1+c_2\psi_2$, where ψ_1 and ψ_2 are represented by two such series, is in general represented asymptotically by the dominant one of the series for $c_1\psi_1$ or $c_2\psi_2$.

If we substitute in the hypothetical series for ψ and equate the term independent of λ to zero, we obtain an equation

$$(5) \qquad P\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right) = 0$$

on removal of a factor $e^{\lambda S}v_0$. Here the explicit expression for P is

$$(6) \qquad P \equiv \xi + \sum_i \xi_i \frac{\partial S}{\partial x_i} + \sum_{ij} \xi_{ij} \frac{\partial S}{\partial x_i} \frac{\partial S}{\partial x_j} + \cdots.$$

Thus P is a polynomial of degree m in $\partial S/\partial x_i$, $(i=1, \cdots, n)$, in which one or more terms of degree m are actually present.

We shall term (5) the "multiplier equation" for $L(\psi, \lambda)=0$. This is an equation of the first order in S which does not contain S, and which is in general non-linear. In the two examples above the multiplier equation is

$$1 + \left(\frac{dS}{dx}\right)^2 = 0,$$

so that we find $S = \pm ix$, up to an additive constant.

It is also to be noted that for a given polynomial P there is one and only one corresponding principal part L_0, while L_1, $L_2, \cdots$ remain entirely arbitrary. For convenience we shall term the special case in which $L \equiv L_0$ the "principal equation" for the given multiplier equation.

Let us proceed to the determination of $v_0, v_1, \cdots$, which turn out to be respectively determined by the later equations for $k=1, 2, \cdots$. We have then to substitute the expressions for ψ, $\partial^{[1]}\psi/\partial x_i, \cdots$ in the power series for $L(\psi, \lambda)$, and equate the coefficients of $1/\lambda$, $1/\lambda^2, \cdots$ to zero. In the case $k=1$, we observe first that a single term only, $Qe^{\lambda S}v_0/\lambda$, is contributed by the terms after the first in the series for $L(\psi, \lambda)$; here

$$Q = \xi^{(1)} + \sum_i \xi_i^{(1)} \frac{\partial S}{\partial x_i} + \cdots.$$

Thus Q is related to L_1 just as P is to L_0. It remains then to determine the term in $1/\lambda$ which arises from L_0. To do so we note that, to terms of the first order in $1/\lambda$,

$$
\begin{aligned}
\frac{\partial^{[1]}\psi}{\partial x_i} &\sim e^{\lambda S}\left(\frac{\partial S}{\partial x_i}v_0 + \frac{1}{\lambda}\left(\frac{\partial v_0}{\partial x_i} + \frac{\partial S}{\partial x_i}v_1\right)\right)\\
&\sim e^{\lambda S}\left(\left[\frac{\partial S}{\partial x_i} + \frac{1}{\lambda}\frac{\partial}{\partial x_i}\right]v_0 + \frac{1}{\lambda}\frac{\partial S}{\partial x_i}v_1\right),\\
\frac{\partial^{[2]}\psi}{\partial x_i \partial x_j} &\sim e^{\lambda S}\left(\left[\frac{\partial S}{\partial x_i} + \frac{1}{\lambda}\frac{\partial}{\partial x_i}\right]\left[\frac{\partial S}{\partial x_j} + \frac{1}{\lambda}\frac{\partial}{\partial x_j}\right]v_0\right.\\
&\qquad\left. + \frac{1}{\lambda}\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}v_1\right),
\end{aligned}
\tag{7}
$$

and so on. Here we use an obvious operational notation.

Now we observe that on substitution in L_0, the terms in v_1 disappear identically because of the multiplier equation. Hence aside from the factor $e^{\lambda S}$, the coefficient of $1/\lambda$ in L_0 is

$$
\begin{aligned}
&\sum_i \xi_i \frac{\partial v_0}{\partial x_i} + \sum_{ij}\xi_{ij}\left(\frac{\partial S}{\partial x_i}\frac{\partial v_0}{\partial x_j} + \frac{\partial S}{\partial x_j}\frac{\partial v_0}{\partial x_i} + \frac{\partial^2 S}{\partial x_i \partial x_j}v_0\right)\\
&+ \sum_{ijk}\xi_{ijk}\left(\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}\frac{\partial v_0}{\partial x_k} + \frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_k}\frac{\partial v_0}{\partial x_j} + \frac{\partial S}{\partial x_j}\frac{\partial S}{\partial x_k}\frac{\partial v_0}{\partial x_i}\right.\\
&\left. + \left(\frac{\partial S}{\partial x_i}\frac{\partial^2 S}{\partial x_j \partial x_k} + \frac{\partial S}{\partial x_j}\frac{\partial^2 S}{\partial x_i \partial x_k} + \frac{\partial S}{\partial x_k}\frac{\partial^2 S}{\partial x_i \partial x_j}\right)v_0\right) + \cdots,
\end{aligned}
$$

where the general law of formation is obvious. But on account of the symmetry relations (4), this may be written

$$
\begin{aligned}
&\sum_i \xi_i \frac{\partial v_0}{\partial x_i} + 2\sum_{ij}\xi_{ij}\frac{\partial S}{\partial x_i}\frac{\partial v_0}{\partial x_j} + 3\sum_{ijk}\xi_{ijk}\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}\frac{\partial v_0}{\partial x_k} + \cdots\\
&\quad + \frac{1}{2}\left(2\cdot 1\sum_{ij}\xi_{ij}\frac{\partial^2 S}{\partial x_i \partial x_j} + 3\cdot 2\sum_{ijk}\frac{\partial S}{\partial x_i}\frac{\partial^2 S}{\partial x_j \partial x_k} + \cdots\right)v_0.
\end{aligned}
$$

But the coefficient of $\partial v_0/\partial x_i$ in the first line is $\partial P/\partial y_i$ if we write $y_i = \partial S/\partial x_i$ in P; and the coefficient of $\partial^2 S/\partial x_i \partial x_j$ in the second line is $(v_0/2)\partial^2 P/\partial y_i \partial y_j$. Hence the required condition for $k=1$ may be written in the form

$$
\sum_i \frac{\partial P}{\partial y_i}\frac{\partial v_0}{\partial x_i} + \frac{1}{2}\left(\sum_{ij}\frac{\partial^2 P}{\partial y_i \partial y_j}\frac{\partial^2 S}{\partial x_i \partial x_j} + Q\right)v_0 = 0. \tag{8}
$$

Let us next throw this linear differential equation in v_0 into a different form, by use of the curves $x_i = x_i(\tau)$, $(i=1, \cdots, n)$, defined by the n ordinary differential equations of the first order

$$\frac{dx_i}{d\tau} = \frac{\partial P}{\partial y_i}, \qquad (i = 1, \cdots, n). \tag{9}$$

Along any such curve in n-dimensional $(x_1 \cdots x_n)$-space, (8) may be written

$$\frac{dv_0}{d\tau} + \Phi v_0 = 0, \tag{10}$$

where Φ is the coefficient of v_0 in (8). The solution of (8) is therefore

$$v_0 = v_0{}^* e^{-\int \Phi d\tau}. \tag{11}$$

Thus the equation for $k=1$ may be looked upon as determining the value of v_0 throughout a tube of these integral curves, once v_0 has been assigned values $v_0{}^*$ on a particular transversal surface Σ. The later coefficients do not enter at this first stage $k=1$.

If now we turn to the later equations for any k $(k>1)$, it is clear that these have a similar form

$$\frac{dv_{k-1}}{d\tau} + \Phi v_{k-1} + A_{k-1} = 0, \tag{12}$$

where A_{k-1} is a known linear differential expression in v_0, $v_1, \cdots, v_{k-2}$. Hence we find that $v_0, v_1, \cdots$ are determined in succession by their values on the transversal surface Σ.

In particular we may suppose that $v_0, v_1, \cdots$ are given arbitrarily on a small region σ of Σ, continuous together with all of their partial derivatives in σ but vanishing along the boundary of and outside of σ. Evidently these functions will then vanish similarly all along the tube and outside of it. We shall refer to a formal solution ψ of this nature, as an "asymptotic wave packet" solution for obvious reasons.

3. *The Associated Canonical Equations.* Let us assume now that S is not only a solution of the partial differential equation $P=0$, but belongs to an n-parameter family of solutions

$$S(x_1, \cdots, x_n; c_1, \cdots, c_{n-1}) + c_n,$$

involving n constants $c_1, \cdots, c_n$, one of which, c_n, is additive. We shall assume that the $n-1$ constants $c_1, \cdots, c_{n-1}$ are "independent" in the sense that if we write $y_i = \partial S/\partial x_i$ so that always, by (6), $P(x_1, \cdots, x_n; y_1, \cdots, y_n) = 0$, the y_i's may be regarded as independent except for the relation just written, that is, we shall assume that the $n \times (n-1)$ matrix

$$||\partial^2 S/\partial x_i \partial c_j||, \qquad (i = 1, \cdots, n; j = 1, \cdots, n-1),$$

is of rank $n-1$. In general an arbitrary (non-singular) solution S of (5) can be imbedded in such a "complete solution."

But by differentiation of the equation $y_i = \partial S/\partial x_i$ along a curve yielding a solution of (9) for any particular set of values of $c_1, \cdots, c_{n-1}$, we obtain

$$\frac{dy_i}{d\tau} = \sum_j \frac{\partial^2 S}{\partial x_i \partial x_j} \frac{\partial x_j}{\partial \tau} = \sum \frac{\partial^2 S}{\partial x_i \partial x_j} \frac{\partial P}{\partial y_j}$$

by use of (9). On the other hand, by partial differentiation of the identity (6) as to x_i, we obtain

$$\frac{\partial P}{\partial x_i} + \sum_{j=1}^{n} \frac{\partial P}{\partial y_i} \frac{\partial^2 S}{\partial x_i \partial x_j} \equiv 0.$$

By combination of this equation and the one which precedes it, we conclude that the following equations also obtain:

$$\frac{dy_i}{d\tau} = -\frac{\partial P}{\partial x_i}, \qquad (i = 1, \cdots, n). \tag{13}$$

It will be observed then that the equations (9), (13) yield a canonical system in x_i, y_i, $(i = 1, \cdots, n)$, of the $2n$th order,

$$\frac{dx_i}{d\tau} = \frac{\partial P}{\partial y_i}, \qquad \frac{dy_i}{d\tau} = -\frac{\partial P}{\partial x_i}, \quad (i = 1, \cdots, n), \tag{14}$$

having a principal function P not containing the independent variable τ.

Moreover we have

$$\frac{d}{d\tau}\left(\frac{\partial S}{\partial c_i}\right) = \sum_j \frac{\partial^2 S}{\partial c_i \partial x_j} \frac{dx_j}{d\tau} = \sum_j \frac{\partial^2 S}{\partial c_i \partial x_j} \frac{\partial P}{\partial y_j} = 0,$$

since $\partial P/\partial c_i \equiv 0$. Hence we infer that the solutions of (14) under consideration satisfy the equations

$$d_i = \frac{\partial S}{\partial c_i}, \qquad (i = 1, \cdots, n-1). \tag{15}$$

But $x_1, \cdots, x_n, y_1, \cdots, y_n$ can evidently be taken as any point on the $(2n-1)$-dimensional manifold $P=0$ at $\tau=0$ so that we obtain in this way the general solution $P=0$ of (14) in the form

$$(16)\ y_i = \frac{\partial S}{\partial x_i}, \ (i = 1, \cdots, n); \quad d_i = \frac{\partial S}{\partial c_i}, \ (i = 1, \cdots, n-1),$$

where $S(x_1, \cdots, x_n, c_1, \cdots, c_{n-1}) + c_n$ is a solution of the multiplier equation of the specified generality. It will be observed that there are $2n-2$ arbitrary constants involved, namely $c_1, \cdots, c_{n-1}, d_1, \cdots, d_{n-1}$. Thus (16) defines a $(2n-2)$-parameter family of curves filling up the $(2n-1)$-dimensional manifold $P=0$; the parameter τ is then determined by setting

$$d\tau = \frac{dx_1}{\partial P/\partial y_1} = \cdots = -\frac{dy_n}{\partial P/\partial x_n},$$

and integrating.

It need hardly be remarked that the equations (14) define the Cauchy characteristics of the partial differential equation $P=0$, in the theory of the solution of which the equations (16) play a well known fundamental part. Thus we may state the following result.

As τ varies, an asymptotic wave packet solution of $L(\psi, \lambda)=0$, belonging to a non-singular phase function S and an amplitude function v_0, travels along the corresponding Cauchy characteristic in $x_1 \cdots x_n$ space, defined by the canonical equations (14), *where the initial values of y_i are given by the equations*

$$y_i = \partial S/\partial x_i, \qquad (i=1, \cdots, n).$$

4. *On Certain Integral Invariants.* We propose next to give the condition (8) an alternative integral invariant form. It is well known that an integral such as

$$I = \int_{V(\tau)} G v_0^k \, dx_1 \cdots dx_n \tag{17}$$

will be invariant as τ changes in case a certain divergence vanishes:

$$\sum_i \frac{\partial}{\partial x_i}\left(Gv_0^k \frac{\partial P}{\partial y_i}\right) \equiv 0.$$

Here $V(\tau)$ is a volume in $x_1 \cdots x_n$ space which moves as τ changes in accordance with equations (14), where $y_i = \partial S/\partial x_i$. This yields

$$\sum_i \frac{\partial G}{\partial x_i}\frac{\partial P}{\partial y_i} v_0^k + kG\left(\sum_i \frac{\partial P}{\partial y_i}\frac{\partial v_0}{\partial x_i}\right) v_0^{k-1} + G\left(\sum_i \frac{\partial^2 P}{\partial x_i \partial y_i} + \sum_{ij} \frac{\partial^2 P}{\partial y_i \partial y_j}\frac{\partial^2 S}{\partial x_i \partial x_j}\right) v_0^k \equiv 0,$$

which by virtue of (8) reduces to

$$(18)\quad \sum_i \frac{\partial P}{\partial y_i}\frac{\partial G}{\partial x_i} + \left(\sum_i \frac{\partial^2 P}{\partial x_i \partial y_i} + \left(1 - \frac{k}{2}\right)\sum_{ij} \frac{\partial^2 P}{\partial y_i \partial y_j}\frac{\partial^2 S}{\partial x_i \partial x_j} - kQ\right) G \equiv 0.$$

In this equation $x_1, \cdots, x_n$ are the independent variables since $\partial S/\partial x_i$ is substituted for y_i throughout. It may be noticed here that $G = \text{const.}$ is a solution for $k = 2$, $Q = 0$, provided that

$$(19)\qquad \sum_i \frac{\partial^2 P}{\partial x_i \partial y_i} \equiv 0,$$

regardless of the choice of y_i; we shall have occasion to employ this result later on.

Conversely it is evident that in general if for the given S and k, a function G is a solution of (18), and if $\int_{V(\tau)} Gv_0^k dx_1 \cdots dx_n$ is an integral invariant for any region V, then v_0 must satisfy the equation (8).

We may now announce the following result.

If v_0 is any amplitude function of the non-singular phase function S and if G satisfies the linear partial differential equation (18) *for this S, and some k, then*

$$I = \int_{V(\tau)} Gv_0^k dx_1 \cdots dx_n$$

is an invariant integral for any volume $V(\tau)$, where $x_1, \cdots, x_n$ vary with τ in accordance with (14) (*with $y_i = \partial S/\partial x_i$*). *Conversely, if this integral is invariant for all regions $V(\tau)$, then v_0 is an amplitude function which satisfies the equation* (8).

5. *The Schrödinger Wave Equation.* Suppose now that we take $n+1$ variables $t, x_1, \cdots, x_n$, with

$$P \equiv \frac{\partial S}{\partial t} + H\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right),$$

so that the "multiplier equation" takes the form of the usual Hamiltonian equation. The corresponding principal equation $L_0(\psi) = 0$ is then the usual Schrödinger wave equation

$$\text{(20)} \quad \frac{2\pi i}{h}\frac{\partial \psi}{\partial t} + H\left(x_1, \cdots, x_n; \frac{2\pi i}{h}\frac{\partial}{\partial x_1}, \cdots, \frac{2\pi i}{h}\frac{\partial}{\partial x_i}\right)\psi = 0,$$

provided we take $\lambda = 2\pi i/h$.

The Schrödinger equation is therefore merely the principal equation which has the usual Hamilton-Jacobi partial differential equation as its multiplier equation, with $\lambda = 2\pi i/h$.

Furthermore, the corresponding Hamiltonian equations are

$$\frac{dt}{d\tau} = 1, \qquad \frac{d\left(\dfrac{\partial S}{\partial t}\right)}{d\tau} = 0,$$

together with

$$\text{(21)} \qquad \frac{dx_i}{d\tau} = \frac{\partial H}{\partial y_i}, \quad \frac{dy_i}{d\tau} = -\frac{\partial H}{\partial x_i}, \quad (i = 1, \cdots, n),$$

with $y_i = \partial S/\partial x_i$. Hence we find $\tau = t + \text{const.}$, $\partial S/\partial t = \text{const.}$ along any trajectory. But in this case a complete solution can be found in the form

$$S = S^*(x_1, \cdots, x_n; c_1, \cdots, c_{n-1}) + c_n t + c_{n+1}$$

with $y_i = \partial S^*/\partial x_i$ for $i = 1, \cdots, n$. Hence the equations for the Cauchy characteristics reduce to the ordinary Hamiltonian equations(21) with $t = \tau$, associated with Schrödinger's wave equation.

Furthermore, it is easily proved that

$$I \equiv \int_{V(\tau)} \psi\bar{\psi}\, dt\, dx_1 \cdots dx_n,$$

where $\bar{\psi}$ denotes the conjugate of ψ, is then an integral invariant to the first order in $1/\lambda$, at least if

$$\sum_i \frac{\partial^2 H}{\partial x_i \partial y_i} \equiv 0. \tag{19'}$$

In fact, since $\lambda = 2\pi i/h$ is a pure imaginary, we have

$$\psi \sim e^{\lambda S} v_0, \quad \bar{\psi} \sim e^{-\lambda S} v_0$$

for S and v_0 real, so that the above integral reduces essentially to $\int_{V(\tau)} v_0^2 dt\, dx_1 \cdots dx_n$, which is of the form treated above with $G \equiv 1$, $k = 2$; furthermore, we have $Q \equiv 0$ in this case. But when rectangular coordinates are employed, we have also

$$H = V(x_1, \cdots, x_n) + \frac{1}{2} \sum k_i y_i^2,$$

where V is the potential energy and $y_1, \cdots, y_n$ are the momental coordinates corresponding to $x_1, \cdots, x_n$, respectively. Hence (19) and (19′) obtain, and the integral I is invariant as stated. Finally, since $d\tau = dt$, it follows at once that $\int_{V(\tau)} v_0^2 dx_1 \cdots dx_n$ also remains constant over any region in $x_1 \cdots x_n$ space. Thus we arrive at the following conclusion.

In the special case of the Schrödinger equation in ψ with rectangular coordinates, the asymptotic wave packets follow the corresponding dynamical trajectories, while the squared amplitude integral $\int |\psi|^2 dx_1 \cdots dx_n$ remains constant over any part of the packet, to terms of the order of h.

6. *On Change of Independent Variables.* Suppose now that we make any change of independent variables

$$\bar{x}_i = f_i(x_1, \cdots, x_n), \qquad (i = 1, \cdots, n).$$

Because of the identities

$$\frac{\partial^{[1]}\psi}{\partial x_i} = \sum_j \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial^{[2]}\psi}{\partial \bar{x}_j},$$

$$\frac{\partial^{[2]}\psi}{\partial x_i \partial x_j} = \sum_{k,l} \frac{\partial \bar{x}_k}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_j} \frac{\partial^{[2]}\psi}{\partial \bar{x}_k \partial \bar{x}_l} + \frac{1}{\lambda} \sum_k \frac{\partial^2 \bar{x}_k}{\partial x_i \partial x_j} \frac{\partial^{[1]}\psi}{\partial \bar{x}_k} + \cdots,$$

the components L_0, L_1, $\cdots$ in L do *not* remain individually invariant in the equation (1), and in particular the principal part L_0 will not carry over into the new principal part by the ordinary rules. In fact the coefficients transform by the rules valid for the attached Hamilton-Jacobi equation. Hence Schrödinger's wave equation in the form (20) is only maintained (in general) under a linear transformation of the independent variables.

This fact indicates that any coordinate system from which we start is to be regarded as a privileged absolute system of reference for the Schrödinger wave equation, up to an arbitrary linear transformation.

7. *Linear Systems and the Dirac Equations.* Let us now turn briefly to a system of k homogeneous linear partial differential equations in $\psi_1, \cdots, \psi_k$:

$$(22)\quad L_i(\psi_1, \cdots, \psi_k; \lambda) \equiv L_{i0} + \frac{L_{i1}}{\lambda} + \cdots = 0, \quad (i = 1, \cdots, k).$$

Here $x_1, \cdots, x_n$ are the independent variables, and the same operational symbols have been introduced as above. Now each L_{i0} may be written as a sum, $\sum_j L_{ij0}$, where L_{ij0} contains only the terms of L_{i0} which involve ψ_j. Furthermore there is then a corresponding set of polynomials $P_{ij}\,(x_1, \cdots, x_n; y_1, \cdots, y_n)$ obtained as in the special case $k=1$ treated above.

If we use formal series solutions,

$$\psi_i = e^{\lambda S}\left(v_{i0} + \frac{v_{i1}}{\lambda} + \cdots\right), \quad (i = 1, \cdots, k),$$

and substitute in the k given equations, the leading terms give us the k equations

$$\sum_j P_{ij}\left(x_1, \cdots, x_n; \frac{\partial S}{\partial x_1}, \cdots, \frac{\partial S}{\partial x_n}\right) v_j = 0, \quad (i = 1, \cdots, k).$$

In order that these be consistent, the "multiplier equation" for the system

$$P \equiv \left| P_{ij} \right| = 0$$

must be fulfilled. Furthermore if the "phase function" S satisfies this multiplier equation, then the linear equations just written determine the k functions ψ_i up to a proportionality factor.

For the determination of this proportionality factor v and so of v_{i0}, of $v_{i1}, \cdots$, we might proceed as before with a greater degree of algebraic complication of course. It is sufficient for our purposes, however, to observe that here too the phenomenon of wave packets occurs. In fact by elimination we may reduce the given system in various ways to a single linear differential equation in a single unknown function ψ, linear in ψ_i and their partial derivatives. Its multiplier equation is then essentially $P=0$, *with the same polynomial* P as before, since the phase functions are the same as before. Hence asymptotic series solutions for ψ having the nature of wave packets exist, associated with this particular P, and so there exist also the corresponding solutions $\psi_1, \cdots, \psi_k$. Thus there exist asymptotic wave packets for the system which follow the Cauchy characteristics belonging to P. The arbitrary proportionality factor in v_{i0} on a transversal surface Σ corresponds to the arbitrary v_0 on Σ in the series for ψ.

Thus it is clear that any *system* of "wave equations" is correlated with a multiplier equation and the allied set of characteristics.

Now the well known work of Sommerfeld showed that the program of Schrödinger leads to a successful theory of the fine structure spectral lines, if one takes account of the special theory of relativity in a natural way; but that it fails to account for certain magnetic properties of the atom. Pauli then substituted a system of two wave equations of the first order for the single Schrödinger equation of the second order so as to bring about the indicated modifications; the multiplier equation $P=0$ obtained is again that necessitated by the special theory of relativity. Finally Dirac obtained a system of four equations of the first order, with multiplier equation $P^2=0$.

Without attempting to analyze the formation of the elegant equations of Dirac, it may be pointed out that the retention of the multiplier equation in unaltered form is of itself sufficient to ensure the proper general form of the characteristic numbers and functions (see the second part of this lecture). It becomes then a very puzzling problem to discover whether the equations of Dirac are to be regarded as more than a set of equations built ad hoc. This is an issue upon which I do not feel myself competent to pronounce.

Part II. The Wentzel-Brillouin-Kramers Method and Asymptotic Series

8. *Formulation of the Problem.*† In order to simplify the form of statement of the problem, we shall write the wave equation to be considered in the form

$$\frac{d^2\psi}{dx^2} + \lambda^2(E - V(x))\psi = 0, \qquad \left(\lambda^2 = \frac{8\pi^2 m}{h^2}\right), \tag{23}$$

and assume that $V(x)$ is real and analytic for all real values of x, and that it possesses a single (absolute) minimum for $x = x_0$, such that $dV/dx = 0$, $d^2V/dx^2 > 0$ at x_0, while $dV/dx \neq 0$ for $x \neq x_0$. Finally we assume that for $|x|$ large, V admits a convergent series expansion in $1/x$ of the form $a + b/x + \cdots$, so that $\lim V(x) = a$ as x becomes infinite. While this is a somewhat idealized form of the case of physical importance, it will be found that the method proposed in justification of the final formula can be extended without essential modification to the cases of physical interest.

We are concerned with the real solutions ψ of (23) which vanish both at $x = -\infty$ and $x = +\infty$, for real and positive λ. Now, for $E \leqq V_0$, the coefficient of ψ in (23) is everywhere negative or 0, and elementary oscillation theorems show that no solution ψ can vanish for $x = \pm\infty$. On the other hand, if $E > a$, this coefficient is everywhere negative, and all solutions oscillate indefinitely often with an amplitude that need not approach 0; this corresponds to the possibility of a continuous spectrum. Hence we may assume that E exceeds V_0 but is less than a.

Let us for the present consider λ as a large positive parameter while E is taken to be restricted as stated. We have then a boundary value problem of classical type, but with the difficulty arising from the singular nature of the boundary conditions (the boundaries lie at infinity) and from the fact that the coefficient of ψ changes sign twice. As far as I know problems of this singular type have not as yet been treated (see, however, Langer, loc. cit.).

† We follow A. Zwaan and J. L. Dunham (loc. cit.) in using a complex variable x. Zwaan's treatment is extremely suggestive, although lacking in essential respects.

9. *An Auxiliary Lemma.* In order to proceed further we shall need the following lemma.

LEMMA. *In any leaf-shaped region σ of the complex x-plane in which $V(x)$ is analytic and which can be covered by a regular family of curves from two points P and Q of its boundary, in such wise that*

$$\mathfrak{R}((V(x) - E)^{1/2}dx) \neq 0 \tag{24}$$

along each curve,† *there will exist two solutions $\psi_i(x, \lambda)$, $(i=1, 2)$, analytic in x and λ, and asymptotically represented by the usual formal series solutions $s_i(x, \lambda)$, $(i=1, 2)$, throughout σ.*

This lemma is a special case of an obvious extension of results contained in my doctoral thesis.‡

10. *Application of the Lemma.* By use of the above lemma, it is possible to determine such regions σ. We restrict attention to the neighborhood of the axis of reals in the x-plane, and let $x=\alpha$ and $x=\beta$ $(\alpha<x_0<\beta)$ denote the two values of x for which $V(x)-E$ vanishes, so that $V(x)-E$ is positive or negative according as x lies outside of or within (α, β). For $x=\alpha$ this function decreases, while for $x=\beta$ it increases.

Let us make a cut in the complex x-plane from $-\infty$ to α, and from β to $+\infty$, along the axis of reals and consider $(V(x)-E)^{1/2}$ in the cut plane; here we take the positive branch on the upper side of the cut $(-\infty, \alpha)$, and then determine this function throughout the cut plane. Evidently $(V(x)-E)^{1/2}$ is a pure imaginary quantity with negative coefficient of i on the real axis between α and β, and is a negative real quantity on the upper side of the cut $(\beta, +\infty)$. On the lower side of the two cuts, the function is of course equal to the negative of its value at the same point on the upper side.

† $\mathfrak{R}$ denotes the "real part of"; E is regarded as fixed.

‡ *On the asymptotic character of the solutions of certain differential equations containing a parameter*, Transactions of this Society, vol. 9 (1908), pp. 219–230. It will be found that for an equation of order n with real parameter ρ (notation of my paper) it is sufficient that along the curves PQ in σ we have

$$\mathfrak{R}(w_1dx) \geqq \mathfrak{R}(w_2dx) \geqq \cdots \geqq \mathfrak{R}(w_ndx).$$

This convenient condition for maintenance of asymptotic form in σ was known to me in 1908, and is indeed obvious from my paper.

Now let us consider the curves

$$\Re((V(x) - E)^{1/2}dx) = 0, \tag{25}$$

which evidently play an important part in the determination of possible regions σ. Under the assumptions made above, the general nature of these curves near the x axis is readily seen to be that indicated in the figure below.

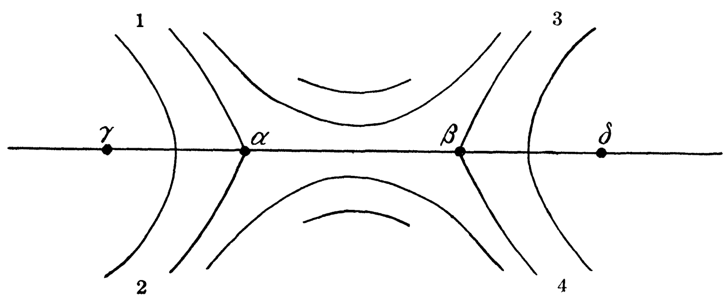

FIG. 1

This leads at once to five special types of regions σ: I: σ above $2\alpha\beta3$; II: σ above $1\alpha\beta4$; III: σ below $1\alpha\beta4$; IV: σ below $2\alpha\beta3$; V: σ between $1\alpha2$ and $3\beta4$.

Here the boundaries of these regions are usually to be excluded. *We shall, however, assume that the point P or Q of the Lemma in regions σ of types* I–IV *may be taken at an infinite end of the real axis.* A critical investigation of the validity of this reasonable assumption is being undertaken by Mr. A. C. Galbraith at Harvard University.

11. *The Distribution of the Characteristic Values.* With these preliminaries in hand we are prepared to determine the distribution of the characteristic values λ. In the first place we are only interested in the solutions ψ of (23) which tend to 0 as x tends to $\pm\infty$. Now we have two formal solutions

$$s_1(x, \lambda) \sim e^{\lambda\int^x (V-E)^{1/2}dx} t_1(x, \lambda), \qquad s_2(x, \lambda) \sim e^{\lambda\int^x (V-E)^{1/2}dx} t_2(x, \lambda),$$

where we take x on the upper side of the cut $(-\infty, \alpha)$ to begin with, and $(V-E)^{1/2}$ as positive there. Here the $t_i(x, \lambda)$ are ordinary power series in $1/\lambda$, and we may suppose that s_1 and t_1 go into s_2 and t_2, respectively, as we traverse the cuts.

Now the unique formal solution of the first type above which reduces formally to 1 for $x=\gamma$ (see the figure) is clearly

$$s_1(x, \lambda)/s_1(\gamma, \lambda) = e^{\lambda\int_\gamma^x (V-E)^{1/2}dx} t_1(x, \lambda)/t_1(\gamma, \lambda).$$

According to the Lemma there is a corresponding ψ_1 having this asymptotic form in the region I, which will evidently approach 0 exponentially as x approaches $-\infty$. Similarly there will exist a solution ψ_2 represented by $s_2(x, \lambda)/s_2(\gamma, \lambda)$ which approaches ∞ under the same circumstances. Hence $\psi_1(x, \lambda)$ is essentially the only solution (up to a constant multiplier) which remains finite as x approaches $-\infty$, and $\psi_1(x, \lambda)/\psi_1(\gamma, \lambda)$ is a special solution, $\psi^*(x, \lambda)$, with the same asymptotic form, which reduces to 1 for $x=\gamma$. This solution is clearly real for x real, since $V(x)$ is real for x real. Hence ψ^* is represented asymptotically by $s_1(x, \lambda)/s_1(\gamma, \lambda)$ in the combined region I+III, since for x below the real axis, ψ^* is conjugate to its value at the conjugate point above the axis, and $s_1(x, \lambda)/s_1(\gamma, \lambda)$ has the same formal property. Taking account of the cut, however, we have

$$\psi_1^*(x, \lambda) \sim \begin{cases} s_1(x, \lambda)/s_1(\gamma, \lambda) \text{ in I, above,} \\ s_2(x, \lambda)/s_2(\gamma, \lambda) \text{ in III, below.} \end{cases}$$

Similarly in the regions II+IV we are led to fix attention upon a solution

$$\psi_2^*(x, \lambda) \sim \begin{cases} s_1(x, \lambda)/s_1(\delta, \lambda) \text{ in II, above,} \\ s_2(x, \lambda)/s_2(\delta, \lambda) \text{ in IV, below,} \end{cases}$$

as yielding the only possible solution which approaches 0 as x becomes positively infinite. It is to be noted that $s_1(x, \lambda)$ is representative of the formal solution which is asymptotic to 0 as x approaches $+\infty$ along the upper side of (β, ∞). Moreover, for a characteristic value and only then, these two solutions must be proportional; that is, their ratio must reduce to essentially the same function $f(\lambda)$ in the overlapping parts above $1\alpha\beta 3$, and below $2\alpha\beta 4$. Hence we are led to the necessary relation

$$\frac{s_1(\delta, \lambda)}{s_1(\gamma, \lambda)} \sim \frac{s_2(\delta, \lambda)}{s_2(\gamma, \lambda)}. \tag{26}$$

Consequently if we write

$$s_i(x, \lambda) = e^{\int^x (d\log\psi_i/dx)dx}, \qquad (i = 1, 2),$$

so that ψ_i is a formal solution of the wave equation (23), the above relation yields

$$\int_\gamma^\delta \frac{d \log \psi_1(x)}{dx} dx + \int_\gamma^\delta \frac{d \log \psi_2}{dx} dx \sim 2k\pi i, \qquad (k \text{ an integer}),$$

or, more briefly, since $d \log \psi_1$ changes to $d \log \psi_2$ as we traverse the cut,

$$\oint \frac{d \log \psi_1}{dx} dx = 2k\pi i,$$

where the path of integration is a positive loop around the points α and β. Written out explicitly this gives the series

$$(27) \quad \oint (V(x) - E)^{1/2} dx + \cdots = \frac{(k + \frac{1}{2})h}{(2m)^{1/2}}, \quad (k = 0, 1, \cdots),$$

which is essentially the desired Wentzel-Brillouin-Kramers equation.

We shall not attempt to consider the degree of precision with which this equation leads to satisfactory approximations to the energy levels E_0, E_1, $\cdots$. It may be anticipated, however, that the approximation will be good whenever the terms of the series on the right diminish rapidly; and that an exact result is obtained in case the series involved converges.

HARVARD UNIVERSITY