On the genesis of wave mechanics

May 27, 2025

Abstract

The ideas which led Erwin Schrödinger to his discovery of wave mechanics are exposed in this paper in a pedagogical manner. We hope, that the material presented may be usefully employed in introductory courses of quantum mechanics, which are usually offered in the last year of the undergraduate curriculum.

1 Introduction

The steps taken by Heisenberg to transform the *old quantum mechanics* into a theory without internal contradictions, culminating in the establishment of *matrix quantum mechanics*, are very well described in a textbook by Tomonaga [1] (We have tried to cite easily accessible references, preferably of a didactic nature). However, the same cannot be said about the birth of wave mechanics. Perhaps because in Schrödinger's first work on the subject [2], the equation that bears his name was postulated practically without any justification. Thus, the usual textbooks [3] provide very little motivation, taking, so to speak, Schrödinger's equation out of the vacuum.

This paper aims to fill this gap in a didactic way, easily applicable to an introductory course on quantum mechanics. Obviously, we have nothing new to say. In fact, this article can be considered a detailed commentary on Schrödinger's second paper on wave mechanics[4]. The original paper is actually more detailed than our section **??**.

Some historical background, essential to understanding Schödinger's motivation, is summarized below. As is well known, the application of Newtonian mechanics, summarized by the Hamilton-Jacobi equation [5] to the atomic structure (complemented by classical electrodynamics), could not make sense of the experimental situation [6]. However, by introducing additional postulates (the Bohr-Sommerfeld quantization rules) it was possible to introduce a certain order into the experimental data: by means of high virtuosity, a set of inconsistent rules (the old quantum mechanics) was applied to a large number of phenomena with admirable success. Note that the Hamilton-Jacobi theory lent itself so well to the implementation of this scheme that one might think it had been invented for this purpose. Consequently, it was felt that the new theory should somehow incorporate certain concepts of the Hamilton-Jacobi theory and reduce to it in the limiting case in which Planck's constant, $\hbar = 1,05457182 \times 10^{-34}$ joule-second, could be neglected.

Another important historical fact is the following: in 1828 Hamilton discovered [7, 8] a formal analogy between the movement of a mass point under the action of a potential $V(\mathbf{r})$ and the propagation of light rays in a medium with a refractive index $n(\mathbf{r})$ at the limit of geometric optics. In fact, it was this analogy that inspired Hamilton to develop his version of mechanics. At the time this fact was considered a curiosity, but Schrödinger asked himself the following question: assuming that the analogy is the reflection of a more fundamental unity of nature, we are led to ask the following question: if geometric optics is the limit of physical optics, when the wavelength of light tends to zero, what is the theory that has Newtonian mechanics as its limit? The situation is depicted in figure 1.

Answering this question led Schrödinger to the discovery of quantum mechanics in its wave form. For didactic reasons, we briefly present the Hamilton-Jacobi theory in Section 2, based on Lagrange's formulation of mechanics. In Section 3, we deduce the main laws of geometric optics from Maxwell's equations, and Section 4 contains the exposition of Schrödinger's ideas, the main part of this article. We also include, as a natural extension of the ideas presented, Feynman's formulation of quantum mechanics.

We believe that this article should be understandable to students who have completed the third year of a bachelor's degree in Physics. The references given were, for the most part, chosen with the didactic point of view in mind.

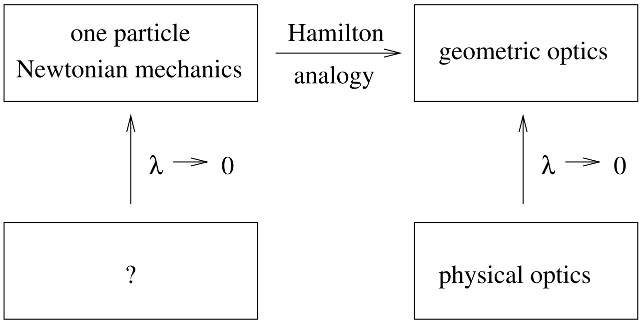

Figure 1: In the limit of small wavelength ($\lambda \to 0$) physical optics goes into geometric optics. In the same limit, what is the theory that goes into classical mechanics, or what is the theory that is the Hamilton analog of physical optics? The answer to this question led Schrödinger to the discovery of the wavelike form of quantum mechanics.

2 Hamilton-Jacobi theory

In this section we introduce some concepts that will be used later [5].

We assume that the Lagrangian formulation of classical mechanics is known, where we specify a system of generalized coordinates $q_1, q_2, \ldots, q_n$. Let $T(q_i, \dot{q}_i)$ be the kinetic energy of the system as a function of the generalized coordinates q_i and the generalized velocities $\dot{q}_i = dq_i/dt$; $i = 1, 2, \ldots, n$. Let us assume that the forces can be derived from a potential $V(q_i)$, such that the total energy of the system is conserved. Newton's equations are equivalent to Lagrange's equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, \ldots, n, \tag{1}$$

where $L(q_1, q_2, \ldots q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n) = T(q_i, \dot{q}_i) - V(q_i)$ is the Lagrangian of the system. For a particle this is obvious, since $L = (1/2)m\mathbf{v}^2 - V(\mathbf{r})$ and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{v}} = \frac{\partial L}{\partial \mathbf{r}} = -\frac{\partial V}{\partial \mathbf{r}}.$$

Since $T(q_i, \dot{q}_j$ is a quadratic function of $\dot{q}_i$, (1) constitutes a system of n equations of second-order in t. Alternatively, we may reduce (1) to a system of $2n$ equations of first order in t by introducing the conjugate momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \qquad i = 1, 2, \ldots, n, \tag{2}$$

and the total energy $H(q_i, p_j)$, expressed in terms of generalized coordinates and momenta

$$H = T + V = 2T - (T - V) = \sum_{i=1}^{n} p_i \dot{q}_i - L. \tag{3}$$

Here, we used

$$2T = \sum_{i=1}^{n} \frac{\partial T}{\partial \dot{q}_i} \dot{q}_i$$

(T is a homogenous function of second order in $\dot{q}_i$!) and $\partial V / \partial \dot{q}_i = 0$ so that

$$2T = \sum_{i=1}^{n} \frac{\partial (T - V)}{\partial \dot{q}_i} \dot{q}_i = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i = \sum_{i=1}^{n} p_i \dot{q}_i. \tag{4}$$

The Hamilton equations are derived calculating [9]

$$\begin{aligned}
\frac{\partial H}{\partial p_i} &\equiv \left(\frac{\partial H}{\partial p_i} \right)_q = \\
&\overset{(3)}{=} \dot{q}_i + \sum_j p_j \left(\frac{\partial \dot{q}_j}{\partial p_i} \right)_q - \left(\frac{\partial L}{\partial p_i} \right)_q \\
&\overset{(2)}{=} \dot{q}_i + \sum_j \left(\frac{\partial L}{\partial \dot{q}_j} \right)_q \left(\frac{\partial \dot{q}_i}{\partial p_i} \right)_q - \sum_j \left(\frac{\partial L}{\partial \dot{q}_j} \right)_q \left(\frac{\partial \dot{q}_i}{\partial p_i} \right) \\
&= \dot{q}_i,
\end{aligned} \tag{5}$$

(the index q indicates that the q's are maintained fixed) and analogously

$$\begin{aligned}
\frac{\partial H}{\partial q_i} &\equiv \left(\frac{\partial H}{\partial q_i} \right)_p \overset{(3)}{=} \sum_j p_j \left(\frac{\partial \dot{q}_j}{\partial q_i} \right)_p - \left(\frac{\partial L}{\partial q_i} \right)_p \\
&= \sum_j p_j \left(\frac{\partial \dot{q}_j}{\partial \dot{q}_i} \right)_p - \left(\frac{\partial L}{\partial q_i} \right)_{\dot{q}} - \sum_j \left(\frac{\partial L}{\partial \dot{q}_j} \right)_q \left(\frac{\partial \dot{q}_j}{\partial q_i} \right)_p \\
&\overset{(1)}{=} \sum_j p_j \left(\frac{\partial \dot{q}_j}{\partial \dot{q}_i} \right)_p - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right)_q - \sum_j \left(\frac{\partial L}{\partial \dot{q}_j} \right)_q \left(\frac{\partial \dot{q}_j}{\partial q_i} \right)_p \\
&\overset{(2)}{=} \sum_j p_j \left(\frac{\partial \dot{q}_j}{\partial \dot{q}_i} \right)_p - \frac{dp_i}{dt} - \sum_j p_j \left(\frac{\partial \dot{q}_j}{\partial q_i} \right)_p =
\end{aligned}$$

$$= -\dot{p}_i. \tag{6}$$

Gathering the results, we obtain the $2n$ equations of first order in t

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \tag{7}$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i. \tag{8}$$

We can also obtain the equations (1), or (7) and (8) as the solutions to variational problems. It is through these that the parallelism between mechanics and geometric optics was historically developed and these results inspired Schrödinger to create his wave mechanics.

We calculate first the variation of $\int \sum_j p_j dq_j$, where we understand a variation δq_i of the generalized coordinates the following: consider the system initially at the point q_{ia} from which it describes a trajectory $q_i(t)$, given by the solution of (1) or (7) and (8): the real trajectory. Consider now a trajectory infinitesimally near (virtual trajectory) given by $\tilde{q}_i(t) = q_i(t) + \delta q_i(t)$. Since the $\delta q_i(t)$ are arbitrary, $\tilde{q}_i(t)$ in general will not fulfill the equations of motion (1) or (7) and (8). We remark that, as dq_i and δq_i are infinitesimals $\delta(dq_i) = d(\delta q_i)$ [10].

Said that, it follows that

$$\begin{aligned} \delta \int \sum_j p_j dq_j &\overset{(2)}{=} \int \sum_j \delta \left(\frac{\partial L}{\partial \dot{q}_j} dq_j \right) \\ &= \int \sum_j \{ \delta \left(\frac{\partial L}{\partial \dot{q}_j} \right) dq_j + \frac{\partial L}{\partial \dot{q}_j} \delta(dq_j) \} \\ &= \int \sum_j \left\{ \delta \left(\frac{\partial L}{\partial \dot{q}_j} \right) \dot{q}_j dt + d \left(\frac{\partial L}{\partial \dot{q}_j} \delta q_j \right) - d \left(\frac{\partial L}{\partial \dot{q}_j} \right) \delta q_j dt \right\} \\ &\overset{(1)}{=} \int \sum_j \left\{ \delta \dot{q}_j dt + d \left(\frac{\partial L}{\partial \dot{q}_j} \delta q_j \right) - \frac{\partial L}{\partial q_j} \delta q_j dt \right\}. \end{aligned} \tag{9}$$

But from (3), we have

$$\delta H = \sum \left\{ \dot{q}_j \delta \left(\frac{\partial L}{\partial \dot{q}_j} \right) + \frac{\partial L}{\partial \dot{q}_j} \delta \dot{q}_j - \frac{\partial L}{\partial q_j} \delta q_j - \frac{\partial L}{\partial \dot{q}_j} \delta \dot{q}_j \right\}$$

$$= \sum \left\{ \dot{q}_j \delta \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} \delta q_j \right\},$$

so that

$$\delta \int \sum_j p_j dq_j = \int \left\{ \sum_j d \left(\frac{\partial L}{\partial \dot{q}_j} \delta q_j \right) + \delta H dt \right\}$$

$$= \sum_j (p_{jf} \delta q_{jf} - p_{ja} \delta q_{ja} + \int \delta H dt, \tag{10}$$

where the indices a and f denote the momenta and the coordinate variations at the initial and the final points of the trajectory, respectively.

If the variations are restricted to vanish at the initial and final instants of time, and if the virtual and real trajectories have the same energy, implying that $\delta q_{ia} = \delta q_{if} = \delta H = 0$, we reach the least action principle

$$\delta \int \sum_{j=1}^{n} p_j dq_j = 0 \quad \text{or} \quad \delta \int 2T dt = 0, \tag{11}$$

for coterminal variations of the same energy. (The second version follows from (4)).

For systems that do not conserve energy (11) can be generalized as follows. Let us consider the variation of the action $\int L dt$ [11]

$$\delta \int L dt = \delta \int (2T - H) dt \overset{(4)}{=} \delta \int (\sum_j p_j \dot{q}_j - H) dt \overset{(10)}{=}$$

$$= \sum_j (p_{jf} \delta q_{jf} - p_{ja} \delta q_{ja}) + \int \delta H dt - \delta \int H dt. \tag{12}$$

The variation $\delta \int H dt$ is due to two terms:
(i) $\int \delta H dt$ due to the variation of energy between the real and virtual trajectories,
(ii) the quantity $(H_f \delta_{ft} - H_a \delta_{at})$ originating from the time variation between the real and virtual trajectories at the end points. As a consequence we have

$$\delta \int L dt = \sum_j (p_{jf} \delta q_{jf} - p_{ja} \delta q_{ja}) + \int \delta H dt - \int \delta H dt$$

$$-(H_f \delta_{ft} - H_a \delta_{at}) =$$

$$= \sum_j (p_{jf}\delta q_{jf} - p_{ja}\delta q_{ja}) - (H_f\delta_{ft} - H_a\delta_{at}). \tag{13}$$

If we subject the virtual trajectory to be co-terminal in space and time, but making no restrictions on energy, we obtain

$$\delta \int Ldt = 0, \tag{14}$$

which is known as Hamilton's principle, which can be stated as follows: a system moves from a configuration to another in such a way that the variation of the integral $\int Ldt$ calculated along the real and virtual trajectories, co-terminal in space and time with the real trajectory, be zero. In other terms, $\int Ldt$ is stationary.

Euler's equations corresponding to the variational problem (14) are exactly the equations (1), which then hold even if the system is not conservative. An advantage of the formulation (14) over (1) is the following: (14) is independent of a particular coordinate system. It follows that equations (1) hold also, whatever the generalized coordinate system chosen.

From the law of action variation (13) we can obtain the derivatives of $S = \int Ldt$ with respect to the initial and final coordinates:

$$\frac{\partial S}{\partial q_{ia}} = -p_{ia}, \tag{15}$$

$$\frac{\partial S}{\partial q_{if}} = p_{if}. \tag{16}$$

If the system is conservative, we have

$$H(q_i, p_j) = E.$$

Replacing herein equations (15) and (16), we get two differential equations, sufficient to determine S (by integration):

$$H(\frac{\partial S}{\partial q_{if}}, q_{if}) = E, \tag{17}$$

$$H(\frac{\partial S}{\partial q_{ia}}, q_{ia}) = E. \tag{18}$$

(Notice that H is quadratic in p_i, so that the negative sign in (15) is irrelevant).

We will not prove here that the integration of equations (17),(18) really determines S: this is a nontrivial problem. Hamilton noted that the determination of S is equivalent to solving the equations of motion. Equations (17) and (18) should be used as follows: we integrate them to obtain S, which is treated as a function of the fixed initial coordinates q_{ia}, of the final variable points q_{if} and of the energy E. By virtue of (15) we can differentiate S with respect to the $(n-1)$ fixed initial coordinates and equate the result to constants (the initial momenta p_{ia}), thus obtaining $(n-1)$ equations in the variables q_{if}. These equations contain $(2n-1)$ constants $(q_{ia}, p_{ia}; i = 1, 2, \ldots, n-1$ and $E)$ and define the trajectory of the system. Notice that, as the energy E is given, we cannot specify arbitrarily all the q_{ia} and p_{ia}, $i = 1, 2, \ldots, n-1, n$, and for that reason we take derivatives only with respect to $(n-1)$ of q_{ia}. Finally, the relations (16) are automatically satisfied because of (18) and allow finding the momenta at any point on the trajectory.

Jacobi developed Hamilton's theory and showed, in particular, that the constants occurring in S need not necessarily be the initial coordinates and that of equations (17) and (18) only one is sufficient to determine S:

$$H(\frac{\partial S}{\partial q_j}, q_j) = E. \tag{19}$$

When the system is not conservative, the total energy is not constant and, instead of energy, time t comes into play. Starting from equation (13) we will introduce the *Hamilton principal function* $W = \int L dt$, where W is considered a function of the final variable points q_{if} and of the time t specifying the upper limit of the integral. Instead of (15) and (16), we now have from (13)

$$\frac{\partial W}{\partial q_{ia}} = -p_{ia}, \quad \frac{\partial W}{\partial q_{if}} = p_{if}, \quad \frac{\partial H}{\partial t} = -E(t), \tag{20}$$

and the analogue of equation (19) becomes

$$H(\frac{\partial W}{\partial q_i}, q_j, t) + \frac{\partial W}{\partial t} = 0. \tag{21}$$

This last one is called the *Hamilton-Jacobi* equation. Jacobi proved that any complete integral of this equation, that is, an integral containing as many arbitrary constants as independent variables (in this case $(n+1)$: t,

q_i; $i = 1, 2, \ldots, n$) can be used. One constant is simply additive, and the complete solution of (21) will be of the form

$$W = f(q_1, q_2, \ldots, q_n; t; \alpha_1, \alpha_2, \ldots, \alpha_n) + A, \tag{22}$$

where A and α_i; $i = 1, 2, \ldots, n$ are constants. To use this complete integral, we prove that the derivatives of W with respect to the constants α_i are also constant, that is,

$$\frac{\partial W}{\partial \alpha_i} = \beta_i, \qquad i = 1, 2, \ldots, n, \tag{23}$$

which is the Jacobi theorem. In fact,

$$\frac{d}{dt}\frac{\partial W}{\partial \alpha_i} = \frac{\partial}{\partial t}\frac{\partial W}{\partial \alpha_i} + \sum_{j=1}^{n} \frac{\partial^2 W}{\partial \alpha_i \partial \alpha_j}\dot{q}_j. \tag{24}$$

On the other hand, if we differentiate (21) with respect to α_i we obtain

$$\frac{\partial}{\partial \alpha_i}\frac{\partial W}{\partial t} + \sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\frac{\partial^2 W}{\partial q_j \partial \alpha_i} = 0, \tag{25}$$

since only $\partial W/\partial q_i = p_i$ depends on α_i. Recalling now Hamilton's equation (8),

$$\dot{q}_j = \frac{\partial H}{\partial p_j},$$

we see that the right-hand side of (24) is equal to the left-hand side of (25), that is,

$$\frac{d}{dt}\frac{\partial W}{\partial \alpha_i} = \frac{d\beta_i}{dt} = 0, \tag{26}$$

and the β_i are indeed constants along the real trajectory.

The momenta are given by

$$p_i = \frac{\partial W}{\partial q_i}, \qquad i = 1, 2, \ldots, n, \tag{27}$$

and we can, through (27), at a certain instant of time t_0, eliminate the n constants α_i in favor of $q_{ia} = q_i(t_0)$. Eliminating β_i from (23),

$$\frac{\partial W(q_{ia}; t_0, \alpha_i)}{\partial \alpha_j} = \beta_j, \tag{28}$$

we obtain the equations of motion as a function of q_{ia}, p_{ia}, by replacing the values of α_i and β_i in (23).

To give some life to this scheme, we calculate the energy levels of the hydrogen atom, according to the procedure of old quantum mechanics. We use the relativistic formulation, but without spin. In this case the total energy is measured from the rest energy

$$E = mc^2(1 - v^2/c^2)^{-1/2} - mc^2 - \frac{e^2}{r}. \tag{29}$$

In planar coordinates r and θ, the velocity v of the electron is obtained by writing the kinetic energy in terms of the momenta p_r and p_θ:

$$\frac{1}{2m}\frac{m^2v^2}{1 - v^2/c^2} = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right).$$

The Hamilton-Jacobi equation becomes

$$mc^2 + \left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 = \frac{1}{c^2}(E + mc^2 + \frac{e^2}{r})^2. \tag{30}$$

Separating the variables

$$S(r, \theta) = S_1(r) + S_2(\theta),$$

we obtain

$$\frac{dS_1}{d\theta} = \alpha_1, \tag{31}$$

$$mc^2 + \left(\frac{\partial S_2}{\partial r}\right)^2 + \frac{\alpha_1}{r^2} = \frac{1}{c^2}(E + mc^2 + \frac{e^2}{r})^2, \tag{32}$$

where α_1 is a constant.

The solution of (31) is

$$S_1 = \alpha_1\theta + \alpha_2, \tag{33}$$

and (32) gives

$$S_2 = \int dr\{-\frac{1}{r^2}(\alpha_1^2 - \frac{e^4}{c^2}) + \frac{2e^2}{rc^2}(E + mc^2) + (\frac{E^2}{c^2} + 2mE)\}^{1/2}. \tag{34}$$

The Bohr-Sommerfeld conditions are introduced by putting

$$\oint p_\theta d\theta = \int_0^{2\pi} \frac{dS_1}{d\theta} d\theta = n_1 h, \tag{35}$$

$$\oint p_r dr = \int \frac{dS_2}{dr} dr = n_2 h. \tag{36}$$

Introducing (33) in (35) eliminate α_i in favor of n_1

$$\alpha_1 = \frac{n_1 h}{2\pi} = n_1 \hbar. \tag{37}$$

To calculate (34) we should find the values of r_0 and r_1 between which r varies when the electron describes its orbit. The integral extends from $r_0 \to r_1 \to r_0$. At the extremes r_0 and r_1, p_r vanishes. Setting the integrand of (34) to zero, to find the roots r_0 and r_1, we obtain the following value for the integral (34)

$$2\pi\{-(\alpha_1^2 - \frac{e^4}{c^2})^{1/2} + \frac{e^2(E + mc^2)}{c^2(-E^2/c^2 - 2mE)^{1/2}},$$

which equating to $n_2 h$ determines E as a function of n_1 and n_2:

$$E = mc^2 \left\{ \left[1 + \frac{\alpha^2}{(n_2 + \sqrt{n_1^2 - a})^2} \right]^{-1/2} - 1 \right\}, \tag{38}$$

where we used (37) and introduced Sommerfeld's fine-structure constant $a = e^2/\hbar c$.

Expanding (38) in powers of $a = e^2/\hbar c$, we obtain the Bohr levels. As we said at the beginning, this calculation shows, by its simplicity (We do not need to calculate trajectories! Try it!), that the Hamilton-Jacobi theory was indeed a very suitable instrument for implementing the rules (35) and (36).

3 Laws of geometric optics [12]

The propagation of electromagnetic waves in a medium described by the refractive index $n(\mathbf{r})$ (in addition to position, it can also depend on the frequency, but we do not make this dependence explicit) is determined by the wave equation

$$\Delta \tilde{f} - \frac{n^2}{c^2} \frac{\partial^2 \tilde{f}}{\partial t^2} = 0, \tag{39}$$

where Δ is the Laplacian

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

and $\tilde{f}$ represents any component of the electric or magnetic field. Assuming that $\tilde{f}$ varies harmonically with time $\tilde{f} = f(x,y,z)e^{-i\omega t}$, we obtain from (39)

$$\Delta f + k^2 f = 0, \qquad k = \frac{\omega n}{c} = \frac{2\pi}{\lambda}, \tag{40}$$

where λ is the wavelength. Geometric optics is the limit when we can neglect the wave character of light and speak only of light rays, that is, when $\lambda \to 0$, or yet, when we can neglect the wavelength compared to the characteristic dimensions of the optical system. The limit $\lambda \to 0$ means $k \to \infty$ and in this case the equation (2) degenerates into $k^2 f = 0$. Despite this, we can extract information from equation (40), casting the solution assuming in the following form

$$f(x,y,z) = A(x,y,z)e^{ik_0\psi(x,y,z)}, \quad k_0 = \frac{\omega}{c} = \frac{2\pi}{\lambda_0}. \tag{41}$$

While f is a function varying rapidly with position (because $k_0 \to \infty$) we may assume that the amplitude A and the eikonal ψ are slowly varying functions of x, y, and z. Inserting (41) into (40), we obtain

$$\Delta f + k^2 f = -k_0 f\left[\left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2 + \left(\frac{\partial\psi}{\partial z}\right)^2 - \frac{k^2}{k_0^2}\right] +$$

$$+2ik_0 f\left[\frac{1}{2}\Delta\psi + \nabla(\ln A)\cdot\nabla\psi\right] + \ldots \tag{42}$$

where the terms indicated by ... do not become infinite when $k_0 \to \infty$. (42) implies that (40) is fulfilled approximately if A and ψ satisfy the equations

$$(\nabla\psi)^2 = n^2(x,y,z), \tag{43}$$

$$\nabla(\ln A)\cdot\nabla\psi = -\frac{1}{2}\Delta\psi, \tag{44}$$

where n is the refraction index. According to equation (41) the surfaces ψ =constant are surfaces of constant phase, that is, the wavefronts are given by ψ =constant, whose normal directions, given by $\nabla\psi$, represent the directions of the light rays. In general $n(x,y,z)$ is not constant and the rays are curved.

In a homogeneous media n=constant, the simplest solution of (43) is

$$\psi = n(\alpha x + \beta y + \gamma z), \quad \text{with} \quad \alpha^2 + \beta^2 + \gamma^2 = 1. \tag{45}$$

The wavefronts are planes and the rays are straight lines parallel to the direction

$$\nabla\psi = n(\alpha, \beta, \gamma). \tag{46}$$

The general solution of (43) is obtained starting from an arbitrary surface and constructing a family of neighboring surfaces (wavefronts), with infinitesimal spacing given by $(|\nabla\psi|)^{-1} = n^{-1}$.

So far we have used only the wave equation (39), which in turn follows from Maxwell's equations, but we can obtain more information using the latter. Let E and $\mathbf{H}$ be fields that vary harmonically

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r})e^{-i\omega t}, \tag{47}$$

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}_0(\mathbf{r})e^{-i\omega t}. \tag{48}$$

In a region free of charges and currents ($\rho = 0$, $\mathbf{j} = 0$), Maxwell's equations are (in Gaussian units)

$$\nabla \times \mathbf{H}_0 + ik_0\varepsilon\mathbf{E}_0 = 0, \tag{49}$$

$$\nabla \times \mathbf{E}_0 - ik_0\mu\mathbf{H}_0 = 0, \tag{50}$$

$$\nabla \cdot (\varepsilon\mathbf{E}_0) = 0, \tag{51}$$

$$\nabla \cdot (\mu\mathbf{H}_0) = 0, \tag{52}$$

Using the Ansatz (41) for $\mathbf{E}_0$ and $\mathbf{H}_0$,

$$\mathbf{E}_0 = \mathbf{e}e^{ik_0\psi}, \tag{53}$$

$$\mathbf{H}_0 = \mathbf{h}e^{ik_0\psi}, \tag{54}$$

in equations (49), (50), (51), and (52), we obtain

$$\nabla\psi \times \mathbf{h} + \varepsilon\mathbf{e} = \frac{i}{k_0}\nabla \times \mathbf{h}, \tag{55}$$

$$\nabla\psi \times \mathbf{e} - \mu\mathbf{h} = \frac{i}{k_0}\nabla \times \mathbf{e}, \tag{56}$$

$$\mathbf{e} \cdot \nabla\psi = \frac{i}{k_0}(\mathbf{e} \cdot \nabla \ln \varepsilon + \nabla \cdot \mathbf{e}), \tag{57}$$

$$\mathbf{h} \cdot \nabla\psi = \frac{i}{k_0}(\mathbf{e} \cdot \nabla \ln \mu + \nabla \cdot \mathbf{h}). \tag{58}$$

In the limit $k_0 \to 0$, we obtain from (55), (56), (57), and (58),

$$\nabla\psi \times \mathbf{h} + \varepsilon\mathbf{e} = 0, \tag{59}$$

$$\nabla\psi \times \mathbf{e} - \mu\mathbf{h} = 0, \tag{60}$$

$$\mathbf{e} \cdot \nabla\psi = 0, \tag{61}$$

$$\mathbf{h} \cdot \nabla\psi = 0. \tag{62}$$

Only equations (59) and (60) are interesting, since (61) and (62) can be obtained from those by scalar multiplication with $\nabla\psi$. The system of equations (59) and (60) for the six unknown quantities $\mathbf{e}$ and $\mathbf{h}$ admits non-trivial solution only if the determinant vanishes. Replacing the value of $\mathbf{h}$ given by (60) in (59), we obtain

$$\frac{1}{\mu}[(\mathbf{e} \cdot \nabla\psi)\nabla\psi - \mathbf{e}(\nabla\psi)^2] + \varepsilon\mathbf{e} = 0 \tag{63}$$

and using (61) results in

$$(\nabla\psi)^2 = n^2(x, y, z) = \mu\varepsilon,$$

which is equation (43).

We can also calculate the time averages of the electric and magnetic densities, $\langle u_e \rangle$ and $\langle u_m \rangle$

$$\langle u_e \rangle = \frac{\varepsilon}{16\pi}\mathbf{e} \cdot \mathbf{e}^*, \quad \langle u_h \rangle = \frac{\mu}{16\pi}\mathbf{h} \cdot \mathbf{h}^*, \tag{64}$$

where $\mathbf{e}^*$ and $\mathbf{h}^*$ are the complex conjugates of $\mathbf{e}$ and $\mathbf{h}$. Using $\mathbf{e}$ and $\mathbf{h}$ from (59) and (60), respectively, results in

$$\langle u_e \rangle = \langle u_h \rangle = \frac{1}{16\pi}(\mathbf{e} \times \mathbf{h}) \cdot \nabla\psi. \tag{65}$$

With (53) and (54) we obtain the time average of the Poynting vector $\langle \mathbf{S} \rangle$ as

$$\langle \mathbf{S} \rangle = \frac{c}{8\pi}\Re(\mathbf{e} \cdot \mathbf{h}^*) = \frac{c}{8\pi\mu}\{(\mathbf{e} \cdot \mathbf{e}^*)\nabla\psi - (\mathbf{e} \cdot \nabla\psi)\mathbf{e}^*\}, \tag{66}$$

or, with (61) and (64),

$$\langle \mathbf{S} \rangle = \frac{2c}{n^2}\langle u_e \rangle\nabla\psi = \frac{c}{n^2}(\langle u_e \rangle + \langle u_m \rangle)\nabla\psi. \tag{67}$$

From (43) it follows that $\nabla\psi/n$ is a unit vector,

$$\mathbf{s} = \frac{\nabla\psi}{n} = \frac{\nabla\psi}{|\nabla\psi|}. \tag{68}$$

Using the velocity $v = c/n$ of the medium (67) becomes

$$\langle\mathbf{S}\rangle = v u_{em}\mathbf{s}, \tag{69}$$

where $u_{em} = \langle u_e\rangle + \langle u_m\rangle$ is the total average energy density. It says, that the the average of the Poynting vector has the direction of the normal to the wavefronts and is equal to the product of the average energy density and the velocity v, i.e. the electromagnetic energy is transported along the rays with velocity v.

Let us also write the ray equation. Let $\mathbf{r}(s)$ be the position vector of a point P on a ray considered as a function of the arc length s, then $\mathbf{s} = d\mathbf{r}/ds$ is a unit vector in the direction of the ray, which in turn is given by $\nabla\psi/n$, that is,

$$n\frac{d\mathbf{r}}{ds} = \nabla\psi, \tag{70}$$

which is the desired equation. With equations (59) and (60) it implies, that the electric and magnetic fields are perpendicular to the rays at all points.

The fundamental equation for rays (70) describes their behavior through the eikonal. We can, however, derive a differential equation governing the propagation of rays using only the refractive index $n(x, y, z)$. Indeed, differentiating (70) with respect to s we obtain:

$$\begin{aligned} \frac{d}{ds}(n\frac{d\mathbf{r}}{ds}) &= \frac{d}{ds}(\nabla\psi) \equiv \frac{d}{ds}(\frac{d\psi}{d\mathbf{r}}) = \\ &= \frac{d\mathbf{r}}{ds}\cdot\frac{d}{d\mathbf{r}}(\frac{d\psi}{d\mathbf{r}}) \equiv (\frac{d\mathbf{r}}{ds}\cdot\nabla)(\nabla\psi) \overset{(70)}{=} \frac{1}{n}(\nabla\psi\cdot\nabla)(\nabla\psi) = \\ &= \frac{1}{2n}\nabla[(\nabla\psi)^2] = \frac{1}{2n}\nabla(n^2), \end{aligned}$$

that is

$$\frac{d}{ds}(n\frac{d\mathbf{r}}{ds}) = \nabla n, \tag{71}$$

which is the vectorial form of the differential equation for the rays. In particular, if the medium is homogeneous, n is constant and (71) yields

$$\frac{d^2\mathbf{r}}{ds^2} = 0, \tag{72}$$

i.e. $\mathbf{r} = \mathbf{a}s + \mathbf{b}$ and rays are straight lines.

The geometrical optics developed so far has been restricted to studying the consequences of equation (43) and covers the area traditionally called geometrical optics, that is, when the wavelength is sufficiently small, the energy transport can be represented by means of a model specified by the scalar function ψ. We could, however, obtain certain geometrical properties of the amplitudes $\mathbf{e}$ and $\mathbf{h}$ from equations (59) and (60). Since, however, polarization properties are not important for understanding the origin of wave mechanics, they will not be considered here.

To emphasize the analogy between geometrical optics and mechanics, we will complement the differential formulation used so far by an integral description by means of variational principles. Let us show that Fermat's principle is equivalent to the ray equation (71). This principle, also known as the *principle of the shortest optical path* (given by $\int nds$), can be stated as follows: the real trajectory between two points P_1 and P_2, followed by a light ray is such that

$$\int_{P_1}^{P_2} nds$$

is stationary, i.e.

$$\delta \int_{P_1}^{P_2} nds = 0. \tag{73}$$

Since

$$ds = [(\frac{dx}{dz})^2 + (\frac{dy}{dz})^2 + 1]^{1/2} dz,$$

the above equation can be written as

$$\delta \int_{P_1}^{P_2} n(x,y,z)[(\frac{dx}{dz})^2 + (\frac{dy}{dz})^2 + 1]^{1/2} dz = 0. \tag{74}$$

To show that (74) implies (71), we recall that Euler's equations, which are necessary conditions for

$$\delta \int F(x,y,z)dz = 0$$

are

$$\frac{\partial F}{\partial x} - \frac{d}{dz}\frac{\partial F}{\partial(dx/dz)} = 0, \quad \frac{\partial F}{\partial y} - \frac{d}{dz}\frac{\partial F}{\partial(dy/dz)} = 0 \tag{75}$$

Applying (75) to the case (74), we obtain

$$\frac{\partial n}{\partial x}[(\frac{dx}{dz})^2+(\frac{dy}{dz})^2+1]^{1/2}$$

$$-\frac{d}{dz}[\frac{n(dx/dz)}{[(dx/dz)^2+(dy/dz)^2+1]^{1/2}}]=0, \tag{76}$$

$$\frac{\partial n}{\partial y}[(\frac{dx}{dz})^2+(\frac{dy}{dz})^2+1]^{1/2}$$

$$-\frac{d}{dz}[\frac{n(dy/dz)}{[(dx/dz)^2+(dy/dz)^2+1]^{1/2}}]=0. \tag{77}$$

That is

$$\frac{d}{ds}(n\frac{dx}{ds})=\frac{\partial n}{\partial x}, \tag{78}$$

$$\frac{d}{ds}(n\frac{dy}{ds})=\frac{\partial n}{\partial y}, \tag{79}$$

while the equation

$$\frac{d}{ds}(n\frac{dz}{ds})=\frac{\partial n}{\partial z}, \tag{80}$$

follows from the identity

$$(\frac{dx}{ds})^2+(\frac{dy}{ds})^2+(\frac{dz}{ds})^2=1.$$

Equations (78), (79), and (80) are identical to equation (71). We may therefore conclude, that the traditional content of geometric optics is equivalent to Fermat's principle. This can yet be formulated in an elegant way by introducing a non-Euclidean space with line element given by

$$d\ell=nds=n(x,y,z)(dx^2+dy^2+dz^2)^{1/2}. \tag{81}$$

Equation (73) is equivalent to saying that light rays propagate along geodesics in the space above. We will see that Schrödinger used a similar non-Euclidean space.

4 Birth of wave mechanics

The geometric optics - mechanics analogy of a particle can be revealed by comparing on the one hand Fermat's principle (73),

$$\delta \int n ds = 0, \tag{82}$$

with the Hamilton's principle for the case in which the energy is conserved (11) (light is monochromatic),

$$\delta \int T dt = \delta \int \frac{m}{2} \mathbf{v}^2 dt = \delta \int \frac{m}{2} |\mathbf{v}|^2 \frac{ds}{|\mathbf{v}|} = \delta \int \frac{m}{2} |\mathbf{v}| ds =$$

$$= \delta \int \sqrt{E - V(x, y, z)}\, ds = 0. \tag{83}$$

From (82) and (83) we conclude that a particle under the action of a potential $V(x, y, z)$ moves in space following the same trajectory of a (monochromatic) light ray in a medium with refraction index $n(x, y, z)$ proportional to $\sqrt{E - V(x, y, z)}$. In order to develop this analogy a little further, we compare the Hamilton-Jacobi equation (19) for a particle,

$$\frac{1}{2m}[(\frac{\partial S}{\partial x})^2 + (\frac{\partial S}{\partial y})^2 + (\frac{\partial S}{\partial z})^2] + V(x, y, z) = E,$$

or

$$(\nabla S)^2 = 2m(E - V), \tag{84}$$

with equation (43)

$$(\nabla \psi)^2 = n^2. \tag{85}$$

Consequently, we identify the eikonal ψ with the action S:

$$\psi \to S \tag{86}$$

$$n^2 \to 2m(E - V) \tag{87}$$

so that the wavefronts correspond to surfaces of constant action S. In Hamilton-Jacobi's theory the momentum is defined by

$$\mathbf{p} = \boldsymbol{\nabla} S, \tag{88}$$

and comparing this equation with the equation (70) for the rays, results in the corespondence

$$n\mathbf{s} = n\frac{d\mathbf{r}}{ds} \to \mathbf{p} \tag{89}$$

While in mechanics we say that the particle propagates perpendicularly to the surfaces of constant action (equation (88)), in geometric optics the rays are orthogonal to the wavefronts. Let us also note that the arrows in equations (86) and (87) cannot be replaced by equalities, since ψ has the dimension $1/k_0 = cT/2\pi$ while S has the dimension of action.

Let us highlight, however, a fundamental difference between the two formulations. While the relationship $v = c/n$ shows that, in optics, the phase velocity is inversely proportional to the refractive index, we conclude from (89) that the particle speed is directly proportional to n. In particular, in vacuum ($n = 1$), the phase speed of the wave fronts is equal to c, while v =constant corresponds to a zero force and the particle trajectory is a straight line with arbitrary speed. (Depending on the initial conditions, it may even be null. In this case, the line degenerates into a point).

So far our optical-mechanical analogy has involved only one particle. It was developed in a more revealing way by Schrödinger [4] using the non-Euclidean space which we alluded to at the end of section III.

Let us consider the general problem of the mechanics of conservative systems, described by the Hamilton-Jacobi equation (21)

$$\frac{\partial W}{\partial t} + T(q_i, \frac{\partial W}{\partial q_j}) + V(q_i) = 0, \tag{90}$$

where W is Hamilton's principal function, the integral $\int (T - V)dt$ along the trajectory (in the configurational space) of the system considered as a function of the final position and time. As the total energy E is constant, we may introduce the action $S(q_i)$ putting

$$W = -Et + S(q_i). \tag{91}$$

Replacing (91) in equation (90), results

$$2T(q_i, \frac{\partial W}{\partial q_i}) = 2(E - V). \tag{92}$$

Since $\partial W/\partial q_i = \partial S/\partial q_i$, we can maintain W in (92).

Let us now consider the configuration space as non-Euclidean [13], which will allow us to extend our analogy to mechanical systems; the correspondence will operate in the configuration space of the system and we will also do geometric optics in this space.

Following Schrödinger [4], we introduce the following non-Euclidean line element

$$ds^2 = 2\bar{T}(q^i, \dot{q}^j)dt^2. \tag{93}$$

Since we will have to carefully distinguish between covariant and contravariant quantities, we write the contravariant vector dq^i with a superscript in accordance with the usual notation. We also indicate by $\bar{T}$ the kinetic energy written in terms of the generalized coordinates and velocities. Since $\bar{T}$ is quadratic in $\dot{q}^i$, (93) is a quadratic form in dq^i. A sufficiently general representation for $\bar{T}$ is

$$\bar{T}(q^i, \dot{q}^j) = \frac{1}{2}\sum_{j=1}^{n} h_j^2(q^i)(\dot{q}^j)^2, \tag{94}$$

so that we can write (93) in the form

$$ds^2 = g_{ij}dq^i dq^j, \tag{95}$$

with

$$g_{ij} = \delta_{ij}h_j^2, \tag{96}$$

$\dot{q}^j$ is a contravariant vector, whereas $p_k = \partial W/\partial q^k$ is a covariant one. The relation between them is

$$p_k = g_{ki}p^i = \delta_{ki}h_k^2\dot{q}^i = h_k^2\dot{q}^k, \tag{97}$$

where we used (96). We can also use (2), that is,

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = \frac{\partial T}{\partial \dot{q}^i} = h_i^2\dot{q}^i.$$

The kinetic energy written in terms of p_k is

$$\sum_{j=1}^{n} h_j^2(\dot{q}^j)^2 = \sum_{j=1}^{n} h_j^2 h_j^{-4} p_j^2 = \sum_{j=1}^{n} h_j^{-2} p_j^2 = 2T(q^i, p_k). \tag{98}$$

The gradient of a scalar function W is a covariant vector of components

$$\nabla_k W = (\frac{1}{h_1}\frac{\partial W}{\partial q^1}, \ldots, \frac{1}{h_n}\frac{\partial W}{\partial q^n}). \tag{99}$$

(Recall that $dW = \nabla_k ds^k$ and that $ds^k = (h_1 dq^1, \ldots, h_n dq^n)$). The square of (99) is

$$(\nabla W)^2 \equiv \nabla_k W \nabla_i W g^{ki}, \tag{100}$$

where g^{ki} is defined by

$$g^{ki} g_{ij} = \delta^k_j. \tag{101}$$

The result is

$$(\nabla W)^2 = \sum_{j=1}^{n} h_j^{-1} p_j^2 = 2T(q^i, p_k) \tag{102}$$

and we can write Hamilton-Jacobi's equation (92) in the following form

$$(\nabla W)^2 = 2(E - V) \tag{103}$$

or

$$|\nabla W| = \sqrt{2(E - V)}. \tag{104}$$

What is the velocity of propagation of the surfaces W=constant? Consider the surface characterized by some value W_0. The surface belonging to $W_0 + dW_0$ is obtainded by erecting normals at each point of the surface W_0 at q^i and advancing along the normal by the distance $ds > 0$. Since $|dW| = |\nabla W| ds = \textit{constant}$ and setting the constant=1 for simplicity, we get

$$ds = \frac{dW_0}{\sqrt{2(E - V)}}.$$

From (91) we see, that in the time span dt, W advances by Edt, so that

$$ds = \frac{Edt}{\sqrt{2(E - V)}}.$$

This yields for surfaces with constant W the phase velocity

$$v = \frac{E}{\sqrt{2(E - V)}}, \tag{105}$$

where the position is encoded by $V(q^k)$. Due to equation $p_k = \partial W/\partial q^k$, the trajectories of the system are orthogonal to surfaces W=constant.

As in the case of one particle, Fermat's principle shows again, that light rays in configuration space propagate along the same trajectories, as long as the refractive index $n(q^i)$ is proportional to $\sqrt{E-V}$. In fact, Hamilton's principle implies Fermat's principle

$$\frac{1}{E}\delta\int_{t_1}^{t_2} 2T dt = \delta\int_{t_1}^{t_2}\frac{2T}{E}dt = \delta\int\frac{\sqrt{2T}}{E}ds =$$

$$= \delta\int\frac{\sqrt{2(E-V)}}{E}ds = \delta\int\frac{ds}{v}, \tag{106}$$

where we used $ds = \sqrt{2T}dt$. Equation (106) is equivalent to Fermat's principle because $1/v = n/c$. Once again, the discrepancy between the phase velocity v (105) and the velocity u of the representative point of the mechanical system in q-space obtained from (93), is evident,

$$u = \frac{ds}{dt} = \sqrt{2T} = \sqrt{2(E-V)}, \tag{107}$$

i.e. $u \sim 1/v$. It is therefore clear that the parallelism between optics and mechanics is not very close. None of the important concepts of optics appear, such as wavelength, amplitude, frequency and wave shape: these are notions that have no analogy in mechanics. W represents the *phase* of the waves, while the meaning of the other notions remains obscure.

At this point Schrödinger observes the following. If all this parallelism serves only to help our intuition, the lack of a closer analogy should not worry us. Classical mechanics in space would have as its analogue a very primitive physical optics. The latter fails when the wavelength of light is not negligible in relation to all the dimensions of the trajectory. We can, however, embrace another point of view, encouraged by the surprising results that the development of wave mechanics provides: *perhaps the failure of geometric optics, in situations in which the wavelength of light cannot be neglected, has as its analogy the failure of classical mechanics in the treatment of atomic phenomena, where electrons follow trajectories with very small dimensions and large curvatures. In this case it will be necessary to look for wave mechanics, and the most natural way to do this is by developing the Hamiltonian image of classical mechanics.*

Let us then hypothesize that the correct development of the preceding analogy leads to *considering the waves of wave mechanics as purely sinusoidal.* This hypothesis is the simplest, but it is arbitrary. The phase of the sine will, of course, be identified with W. The arbitrariness consists in saying that the phase of the sine is a linear function of W. As the sine argument must be dimensionless, while W has an action dimension, let us introduce a universal constant, that is independent not only of the energy, but also of the system considered. We can write this constant in the form $2\pi/h$. The time-dependent factor will then be of the form

$$\sin(\frac{2\pi}{h}W + \text{const}) \overset{(91)}{=} \sin(-\frac{2\pi}{h}Et + \frac{2\pi}{h}S(q^i) + \text{const}), \tag{108}$$

from where we take the *frequency of the waves*,

$$\nu = \frac{E}{h}, \tag{109}$$

in a perfectly natural way. Once we have the frequency, we can determine the wavelength using (105)

$$\lambda = \frac{v}{\nu} = \frac{h}{\sqrt{2(E-V)}}. \tag{110}$$

The relation between the phase velocity and the frequency

$$v = \frac{h\nu}{\sqrt{2(h\nu - V)}} \tag{111}$$

yields the dispersion law of these waves. Hence, following de Broglie we can calculate the group velocity v_g

$$v_g = \frac{d\omega}{dk} = \frac{d\nu}{d(\lambda^{-1})}.$$

From (110), we have

$$\frac{d(\lambda^{-1})}{d\nu} = \frac{d}{d\nu}[\sqrt{2(h\nu - V)}/h] = \frac{1}{\sqrt{2(E-V)}},$$

that is,

$$v_g = \sqrt{2(E-V)}, \tag{112}$$

which is, by comparison with equation(107), exactly the velocity u of the representative point of the system in space q. In the case of one particle, we moreover have

$$u = \frac{ds}{dt} = \sqrt{m}\frac{dq}{dt} = \sqrt{2(E-V)} = \frac{p}{\sqrt{m}}.$$

Comparing with (110) the result is $\lambda_{Non-euclidean} = \sqrt{m}\frac{h}{p}$. Remember that due to (93) $ds = \sqrt{m}dq$, so that reverting to the usual Euclidean metric, we get

$$\lambda = \frac{h}{p}, \tag{113}$$

which is the ratio between *wavelength* and momentum obtained by de Broglie [14]. At this point Schrödinger [4] discusses the construction of a wave packet in space q by superposition of monochromatic plane waves. This discussion can be summarized as follows. Except for the constant h^{-1}, W represents the phase of the wave function. Suppose we have at our disposal not just a single wave system, but a *continuous set*, defined by the variation of certain parameters α_i. In this case, the conditions $\partial W/\partial\alpha_i$=constant mean that the system of neighboring waves are in phase. The equations $\partial W/\partial\alpha_i$=constant then determine the geometric locus of phase concordance, which will be a point for a sufficiently large number of equations. Remembering equations (23), (27), and (28), and identifying the α_i constants with those of the same name in equation (22), we reach the conclusion that the geometric locus where a certain set of systems of waves depending on n parameters is in phase, moves according to the same laws as the point represented in the space q of the corresponding classical system.

Continuing the elaboration of the wave theory, Schrödinger is led to attribute a physical meaning to the propagation of those waves with a concomitant abandonment of mechanical concepts.

It is natural, then, to replace the Hamilton-Jacobi equation by a wave equation in space q. We already know the *phase velocity* of the waves, given by equation (105). The simplest propagation equation that can be written will be

$$\Delta\phi - \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = 0, \tag{114}$$

where Δ is the Laplacian and ϕ a certain *wave function*. For a system with more than one particle ϕ will be a function of at least six variables, e.g., $\mathbf{r}_1$

and $\mathbf{r}_2$ for two particles. This fact caused Schrödinger (and others) a huge headache due to the difficulty of interpreting a wave phenomenon in a space of six or more dimensions. Using (105) in the equation above, we obtain:

$$\Delta\phi - \frac{2(E-V)}{E^2}\frac{\partial^2\phi}{\partial t^2} = 0. \tag{115}$$

But this cannot be the correct equation, for at least two reasons:

i) since the waves are monochromatic (i. e., they depend on time as $e^{2\pi i\nu t}$), we cannot allow the simultaneous appearance of the constant E and time derivatives;

ii) if we consider two isolated systems with energies E_1 and E_2, the energy of the combined system will be $E = E_1 + E_2$. If ϕ_1 and ϕ_2 are the wave functions of the systems 1 and 2, they should fulfill the equation (115) with energies E_1 and E_2, respectively. But then $\phi = \phi_1\phi_2$ does not satisfies the equation (115) with $E = E_1 + E_2$ due to the non-linearity of E in the equation [15].

We then use equation (108) to eliminate the time derivatives from (115) to obtain

$$\frac{\hbar^2}{2}\Delta\phi + (E-V)\phi = 0, \tag{116}$$

which is the famous **Schrödinger equation for stationary processes**. Note that the Laplacian must be computed with the non-Euclidean metric (95). For one particle we obtain

$$\{-\frac{\hbar^2}{2m}\nabla^2 + V\}\phi = E\phi. \tag{117}$$

It seems a trivial matter to extend this equation to non-stationary processes with arbitrary time dependence. Since, in the stationary case, the time dependence is of the type $e^{\pm iEt/\hbar}$, we have

$$E\phi = \pm i\hbar\frac{\partial\phi}{\partial t}. \tag{118}$$

If we eliminate E from equation (116), we obtain (the opposite sign gives the complex conjugate equation)

$$(-\frac{\hbar^2}{2}\Delta + V)\phi(q^i, t) = \imath\hbar\frac{\partial\phi(q^i, t)}{\partial t}, \tag{119}$$

which is the **Schrödinger equation for phenomena with arbitrary time dependence**. However, it took Schrödinger a long time to discover this simple generalization. The problem is that equation (119) is the first equation in physics with imaginary coefficients. Consequently ϕ is necessarily a complex function and cannot describe a real wave phenomenon *accompanying*, for example, an electron in real space. ϕ also satisfies the real equation

$$\{(-\frac{\hbar^2}{2}\Delta + V + \imath\hbar\frac{\partial}{\partial t})(-\frac{\hbar^2}{2}\Delta + V - \imath\hbar\frac{\partial}{\partial t})\phi = 0, \tag{120}$$

which is the *elastic plate* equation, and for some months Schrödinger thought that (120) was the correct equation. He actually also tried the relativistic Klein-Gordon equation, but eventually realized that its non-relativistic limit produced the correct hidrogen spectrum.

The little letter $\imath$ raised a sea of troubles, felt still today after a century of dedicated research to "understand quantum mechanics". I suspect, that this will never come to pass. After all we are classical objects, "understand" being only a vaguely defined concept. But the reader may consult reference [18] to see for himself, how far people are willing to go in order to fit quantum mechanics into the Procrustes bed of a classical observer.

We hope that the patient reader has gained some benefit by getting this far. Before finishing, we would like to present another way of approaching wave mechanics due to R. Feynman [16]. As he said: "I tried to understand quantum mechanics my way".

To do this, we recall the Hamilton-Jacobi equation (103) and its optical analogue, which is equation (43). The method for obtaining a general solution of these equations, described immediately after equations (7) and (8), is nothing more than Huygens' principle in its primitive formulation. Let us recall that the rigorous formulation is Kirchhoff's [17] and only by means of approximate hypotheses about the boundary conditions do we arrive at Huygens' principle. The latter allows the determination of electromagnetic fields at any point in space, from knowledge of the fields on a surface ξ. Since knowledge of second derivatives is not necessary in this construction, propagation is of the diffusion type, while Kirchhoff's formulation is equivalent to a wave equation, involving second derivatives in time. Noting that Schrödinger's equation is of first order in $\partial/\partial t$, we will construct an optical equation in space q using Huygens' primitive principle, which states: consider all points from an initial wavefront $\phi(q_i)$ as secondary wave emitting centers. The new wavefront is the envelope of the secondary waves. We already know that the phase of the waves is given by

$$\frac{\imath}{\hbar}\int_{t_1}^{t_2} L(q,\dot{q})dt.$$

Summing over all points of the initial wavefront with the appropriate phase, we obtain

$$\phi(q_{i2}) = \int \exp\{\frac{\imath}{\hbar}\int_{t_1}^{t_2} L(q,\dot{q})dt\}\phi(q_{i1})dq_{i1} \tag{121}$$

where the integration

$$\int_{t_1}^{t_2} L(q,\dot{q})dt$$

is performed along a real trajectory. The quantum generalization of (121) due to Feynman consists in saying that we must consider in (121) not only the real path that satisfies the classical equations of motion, but we must sum over all paths, connecting (q_{i1}, t_1) and (q_{i2}, t_2), that is,

$$\phi(q_{i2}, t_2) = \frac{1}{N}\sum_{\text{traj.}}\int \exp\{\frac{\imath}{\hbar}\int_{t_1}^{t_2} L(q,\dot{q})dt\}\phi(q_{i1}, t_1)dq_{i1} \tag{122}$$

where N is a normalization factor. Note that, in the classical limit in which $\hbar \to 0$, the exponent will oscillate violently, the contributions to the intgral in dq_{i1} cancelling out, unless the exponent be stationary, that is,

$$\delta\int_{t_1}^{t_2} Ldt = 0,$$

which is equation (14). One of the advantages of Feynman's formulation is the simplicity with which the classical limit is obtained. The definition of the sum (integral) over the trajectories is a mathematical problem that remains open to this day. The equivalence (apart from mathematical technicalities) between the usual formulation and Feynman's is discussed in ref. [16] and also in ref.[19].

It is a pleasure to acknowledge the lively discussions on the subject with my colleagues R. C. T. da Costa, L. G. de Freitas e R. L. Lobo e Silva.

References

[1] S. Tomonaga, *Quantum Mechanics*, vol. 1, North-Holland, Amsterdam, 1968. See also M. Jammer, *The Conceptual Development of Quantum Mechanics*, McGraw-Hill, 1966. This work is quite faithful from a historical point of view, but it was not written for didactic purposes.

[2] E. Schrödinger, Annalen der Physik **79**, 361-376 (1926); received for publication on January 27, 1926.

[3] L. Schiff, *Quantum Mechanics* McGraw-Hill; A. Messiah, *Mécanique Quantique*, Dunot, Paris; L. D. Landau and E. M. Lifschitz, *Quantum Mechanics*, Pergamon.

[4] E. Schrödinger, Annalen der Physik **79**, 489-527 (1926); received for publication on February 23, 1926. In 1933 a French translation of Schrödinger's works appeared (*Mémoires sur la mécanique ondulatoire*, Alcan, Paris, 1933) with commentary by the author.

[5] In addition to the usual textbooks (e. g. H. Goldstein, *Classical Mechanics*, Addison-Wesley, 1959; L. D. Landau and I. M. Lifshitz, *Mechanics*, Pergamon; E. J. Saletan and A. H. Cromer, *Theoretical Mechanics*, Wiley, 1971), we recommend the little book W. Yourgrau and S. Mandelstam *Variational Principles in Dynamics and Quantum Theory*, Pitman and Sons, 3rd edition, 1968.

[6] See, for instance, Max Born, *Atomic Physics*, Blackie and Son, London.

[7] W. R. Hamilton, "Theory of Systems of Rays", Trans. Royal Irish Academy *15*, 69 (1828). A more accessible reference would be ref. [8].

[8] M. Born and E. Wolf, *Principles of Optics*, Pergamon, 1959.

[9] The number of an equation on top of an equality sign means that the equation was used to establish the equality. Thus $A \overset{(1)}{=} B$ means that the equation (1) was used to transform A into B.

[10] This can also be seen as follows: let us parametrize the variations $\delta q_i(t)$ for a parameter (or several parameters) α, such that $q_i(t, \alpha) = q_i(t)$ for

$\alpha = 0$, the real trajectory, and $q_i(t,\alpha) = \tilde{q}_i(t)$ for any α, the virtual trajectory. Then

$$\delta q_i(t,\alpha) = \frac{\partial q_i(t,\alpha)}{\partial \alpha} d\alpha$$

and

$$dq_i(t,\alpha) = \frac{\partial q_i(t,\alpha)}{\partial t} dt d\alpha$$

$$\delta(dq_i) = \delta\left(\frac{\partial q_i}{\partial t} dt\right) = \frac{\partial^2 q_i}{\partial \alpha \partial t} dt d\alpha = \frac{\partial^2 q_i}{\partial t \partial \alpha} d\alpha dt = d(\delta q_i)$$

provided that we use the equality

$$\frac{\partial^2 q_i}{\partial \alpha \partial t} = \frac{\partial^2 q_i}{\partial t \partial \alpha}.$$

Differentiability up to second order will always be assumed as valid.

[11]

$$\int L dt = \int_{t_1}^{t_2} L[q_i(t), \dot{q}_j(t)] dt$$

[12] See for example [8].

[13] This formulation dates back to H. Hertz and F. Klein.

[14] The works of L. de Broglie that inspired Schrödinger can be found in *Selected Papers on Wave Mechanics* by Louis de Broglie and Léon Brillouin, Blackie and Son, 1929.

[15] This observation is due to Einstein (who learned about the work of Schrödinger through Planck) who, initially, thought that Schrödinger had used the equation (115). To illustrate his objection, Einstein wrote an equation that did not suffer from this defect — it was equation (116). See *Briefe zur Wellenmechanik*, Schrödinger, Planck, Einstein, Lorentz, edited by K. Przibram, Springer, Vienna, 1964, p. 21. English version: *Letters on Wave Mechanics*, translated by M. J. Klein, Vision Press, 1967.

[16] See the book by Yourgrau and Mandelstam, ref. [5] and R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, 1965.

[17] See e. g. ref. [7], p. 374.

[18] Ch. Stoica, arXiv:1311.0765v2, 2015; *It From Bit or Bit From It?*, Springer International Publishing, 2015, pages 51-64; R. Chaves, G. Barreto Lemos and J. Pienaar, Phys. Rev. Lett. 120, 190401, 2018.

[19] R. Köberle, *Introduction to path-integral Quantum Field Theory*, Rev.Bras.Ens. de Fisica, **43**(2021).