

A MAXIMUM ENTROPY APPROACH TO WAVE MECHANICS

Juho Leppäkangas¹

ABSTRACT.

We employ the maximum entropy principle, in the context of Bayesian inference by the impersonal physical interactions, together with the experimental Heisenberg's uncertainty principle for a simple construction of the general wave mechanical static state of a single, interacting mass particle with no internal degrees of freedom. Subsequently, this first principle approach allows to derive via Newtonian mechanics the dynamical equation of motion in the realm of non-relativistic wave mechanics, i.e., the Schrödinger equation.

¹Email: juhole@hotmail.com

1. INTRODUCTION

Since the advent of the modern wave mechanics and Schrödinger's seminal papers [1], the origins of the theory have been a subject of much discussion and several reasonable approaches have been proposed over the years, including Ref. [2] where a particular form of the uncertainty principle is assumed in order to derive the Schrödinger equation, that is, the fundamental wave mechanical equation of an epistemic entity called wave function. The previous approach to wave mechanics is quite compelling as it allows a proper extension of classical mechanics based on a one new assumption, originally found in a conceptual form by Heisenberg [3]. Consequently, the uncertainty principle is at the heart of wave mechanics and it allows to isolate the source of the new physical phenomena. In this presentation, that in a sense is analogous with the statistical approach taken in Ref. [2], an experimentally inferred uncertainty principle together with the maximum entropy principle [4] are used to construct the general static wave function. Subsequently, Newtonian mechanics provides additional physical constraints to morph the probabilistic ensemble, hence to obtain the corresponding dynamical equation of motion.

2. THE WAVE MECHANICAL MODEL

The well-known single- or many-slit electron diffraction experiment illustrates a non-classical phenomenon as individual electrons form a diffraction pattern as a collective behaviour [5]. Therefore, electrons, like other point-like particles with small mass, behave statistically: every identical particle is observed at a random location but many observations form a probability distribution that is similar to the optical phenomena where photons form an intensity pattern that can be explained by the Huygens' principle and wave interference. Hence, we can conclude that the motion of every electron is wave-like and therefore it is impossible for an individual electron to travel along a classical path.² However, classically the state of motion of a particle is represented by its position \mathbf{x} and momentum \mathbf{p} , and we can further consider a statistical version of classical mechanics where a state is a probabilistic variable. The wave-like interference pattern implies that the classical statistical mechanics is inadequate, but we may still assume that this peculiar distribution of microstates obeys the universal principle under which the entropy of a physical system tends to be maximized [4]. Here the entropy is understood to be the general measure of uncertainty of a probability distribution and every physical interaction can contribute to the objective, i.e., global, knowledge about the system. We also recall that in a thermodynamical system a pair (\mathbf{x}, \mathbf{p}) corresponds to an energy microstate, and physical interactions dissipate kinetic energy and thus increase entropy in the surroundings. Analogously, the macrostate of an interacting particle, i.e., the (\mathbf{x}, \mathbf{p}) -distribution, has a non-zero entropy.

Let us examine the single-slit electron diffraction experiment where the electrons arrive from far away with a definite momentum. When the slit is made more narrow, so that the location of those electrons that move through the slit become more accurate, we can observe that the direction of motion becomes more uncertain and the momentum spreads. Hence, we may assume that the kinematic system of a single particle is a joint distribution of position and momentum, and the statistical behaviour is understood through a new physical constraint. In particular, we adopt the well-known *Heisenberg's uncertainty principle*: if the position/momentum of a particle becomes more localized, then its momentum/position becomes more spread out [5]. The content of information of the position of a particle around some point corresponds to how much the information of momentum is uncertain. Thus, for every point in the position space we can attach an infinite array with elements representing probability for points in the momentum space, implying the degree of dispersion. In the spirit of the *maximum entropy principle* [4] we can model this by a function $\psi(\mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^3$, that represents an unbiased estimation of the momentum distribution and a convenient representation is a linear weighted sum

²Here we have two basic options: the interference of probabilities or the uncertainty relation of motion.

$$(2.1) \quad \psi(\mathbf{x}) = \sum_{\mathbf{p} \in \mathbb{R}^3} a(\mathbf{p}) \mathbf{u}(\mathbf{p}, \mathbf{x}),$$

where the unspecified vectors $\mathbf{u}(\mathbf{p}, \mathbf{x})$ form an orthonormal basis, suitable inner-product $\langle \cdot, \cdot \rangle_{\mathbf{x}}$ induces a norm and terms $a(\mathbf{p})$ are scalars. We require that $\langle \psi, \psi \rangle_{\mathbf{x}} = \sum_{\mathbf{p}} |a(\mathbf{p})|^2 = 1$ and that $|a(\mathbf{p})|^2$ gives a probability for \mathbf{p} . Now the function $\psi(\mathbf{x})$ represents in a position-dependent way the distribution of momentum. Similarly, we can assume that a change in the information of momentum changes the position distribution, thus we have a function $\phi(\mathbf{p})$, for $\mathbf{p} \in \mathbb{R}^3$, that contains an unbiased estimation of the position distribution:

$$(2.2) \quad \phi(\mathbf{p}) = \sum_{\mathbf{x} \in \mathbb{R}^3} b(\mathbf{x}) \mathbf{v}(\mathbf{x}, \mathbf{p}),$$

where the unspecified vectors $\mathbf{v}(\mathbf{x}, \mathbf{p})$ form an orthonormal basis, suitable inner-product $\langle \cdot, \cdot \rangle_{\mathbf{p}}$ induces a norm and terms $b(\mathbf{x})$ are scalars. Again, we have $\langle \phi, \phi \rangle_{\mathbf{p}} = \sum_{\mathbf{x}} |b(\mathbf{x})|^2 = 1$ and $|b(\mathbf{x})|^2$ gives a probability for \mathbf{x} . Since the both position and momentum are continuous variables, we can conclude that both $a(\mathbf{p})$ and $b(\mathbf{x})$ are continuous functions that behave inversely to each other: if $a(\mathbf{p})$ vanishes outside of a narrow interval $\Delta \mathbf{p}$, then $b(\mathbf{x})$ must be non-zero on a large interval $\Delta \mathbf{x}$ – vice versa. Hence, the state of motion of a wave-like mass particle is represented by the following coupled, continuous and entropy driven fields, here denoted as *wave functions*:

$$(2.3) \quad \psi(\mathbf{x}) = \int_{\mathbb{R}^3} a(\mathbf{p}) \mathbf{u}(\mathbf{p}, \mathbf{x}) d\mathbf{p} \quad \text{and} \quad \phi(\mathbf{p}) = \int_{\mathbb{R}^3} b(\mathbf{x}) \mathbf{v}(\mathbf{x}, \mathbf{p}) d\mathbf{x}.$$

From the previous assumptions and the theory of optical diffraction we may infer that a suitable method to connect the wave functions ψ and ϕ is a unitary Fourier transform, where the transformation pair contains both the spectrum and the intensity of the random variables. Here we note that any vector $f(\mathbf{x})$ in the infinite-dimensional space $L^2(\mathbb{R}^3; \mathbb{C})$ can be projected onto an orthonormal basis made of vectors $\exp(i \langle \mathbf{k}, \mathbf{x} \rangle) / \sqrt{2\pi}^3$ or vectors $\exp(-i \langle \mathbf{k}, \mathbf{x} \rangle) / \sqrt{2\pi}^3$. Since both the position and momentum of a physical particle are finite, both functions $a(\mathbf{p})$ and $b(\mathbf{x})$ must vanish as $|\mathbf{x}|, |\mathbf{p}| \rightarrow \infty$, i.e., low entropy microstates are negligible, and we may assume³ that $\psi, \phi \in L^2(\mathbb{R}^3; \mathbb{C})$. Therefore, the wave function ψ can be represented by its Fourier transform $\hat{\psi}$ as follows, see Ref. [6]:

$$(2.4) \quad \psi(\mathbf{x}) = \frac{1}{\sqrt{2\pi}^3} \int_{\mathbb{R}^3} \hat{\psi}(\mathbf{k}) \exp(i \langle \mathbf{k}, \mathbf{x} \rangle) d\mathbf{k},$$

where the components of \mathbf{k} have the dimension of [1/length]. In addition, the wave function ϕ can be represented by its inverse Fourier transform $\check{\phi}$ as follows:

$$(2.5) \quad \phi(\mathbf{p}) = \frac{1}{\sqrt{2\pi}^3} \int_{\mathbb{R}^3} \check{\phi}(\mathbf{q}) \exp(-i \langle \mathbf{q}, \mathbf{p} \rangle) d\mathbf{q},$$

where the components of \mathbf{q} have the dimension of [1/momentum]. Since ψ and ϕ contain the probability distribution of both position and momentum, respectively, we assume that those distributions correspond to the spectral densities $|\check{\phi}|^2$ and $|\hat{\psi}|^2$. Now, due to the probability conditions and Plancherel's theorem, we have

$$(2.6) \quad \int_{\mathbb{R}^3} |\psi(\mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^3} |\hat{\psi}(\mathbf{k})|^2 d\mathbf{k} = 1 \quad \text{and} \quad \int_{\mathbb{R}^3} |\phi(\mathbf{p})|^2 d\mathbf{p} = \int_{\mathbb{R}^3} |\check{\phi}(\mathbf{q})|^2 d\mathbf{q} = 1.$$

³The choice of complex basis is due to its mathematical simplicity and the fact that we are only interested of intensities, but we can always use the Euler's formula and choose a real two dimensional basis and represent the intensity as a sum of two squares, e.g., $|a_1 + ia_2|^2 = a_1^2 + a_2^2$.

We note that the wave functions ψ and ϕ can be connected if we substitute $\mathbf{p} := C\mathbf{k}$, where C is an experimental constant that has the dimension of $[\text{length} \times \text{momentum}] = [\text{time} \times \text{energy}]$. We also need to substitute $\phi(\mathbf{p}) := \hat{\psi}(\mathbf{k})/\sqrt{C^3}$. The probability condition remains unchanged:

$$\int_{\mathbb{R}^3} |\phi(\mathbf{p})|^2 d\mathbf{p} = \int_{\mathbb{R}^3} \left| \hat{\psi}(\mathbf{k})/\sqrt{C^3} \right|^2 C^3 d\mathbf{k} = \int_{\mathbb{R}^3} \left| \hat{\psi}(\mathbf{k}) \right|^2 d\mathbf{k} = \int_{\mathbb{R}^3} |\psi(\mathbf{x})|^2 d\mathbf{x}.$$

It turns out that the appropriate constant C is the reduced *Planck's constant* \hbar , see Ref. [5]. Along these lines, we have ended up with the following statistical model that describes the intrinsically incomplete information content of a state of motion:

$$(2.7) \quad \psi(\mathbf{x}) = \frac{1}{\sqrt{2\pi\hbar^3}} \int_{\mathbb{R}^3} \phi(\mathbf{p}) \exp(i \langle \mathbf{p}, \mathbf{x} \rangle / \hbar) d\mathbf{p};$$

$$(2.8) \quad \phi(\mathbf{p}) = \frac{1}{\sqrt{2\pi\hbar^3}} \int_{\mathbb{R}^3} \psi(\mathbf{x}) \exp(-i \langle \mathbf{p}, \mathbf{x} \rangle / \hbar) d\mathbf{x},$$

where $|\psi|^2$ and $|\phi|^2$ correspond respectively to position and momentum probability distribution. The previous pair of ψ and ϕ satisfies the desired physical connection: if one changes, then the other one changes as well, but the terms $|\psi|^2$ and $|\phi|^2$ remain as well-defined distributions. Furthermore, due to the Heisenberg's uncertainty principle it is impossible to localize both distributions. To this end, we need to have a suitable model⁴ for the uncertainty:

If f belongs to $L^2(\mathbb{R}^3; \mathbb{C})$ and $\int_{\mathbb{R}^3} |f(\mathbf{x})|^2 d\mathbf{x} = 1$, then the standard deviation of the function f about the point $\mathbf{a} \in \mathbb{R}^3$ is

$$(2.9) \quad \Delta_{\mathbf{a}}(f) := \sqrt{\int_{\mathbb{R}^3} |\mathbf{x} - \mathbf{a}|^2 |f(\mathbf{x})|^2 d\mathbf{x}}.$$

It is clear that the deviation $\Delta_{\mathbf{a}}(f)$ is a measure of how much the function f is localized around the point $\mathbf{a} \in \mathbb{R}^3$, e.g., the term is small if f is non-zero only around the point $\mathbf{a} \in \mathbb{R}^3$. Thus, $\Delta_{\mathbf{x}'}(\psi)$ is the standard deviation of the position distribution and $\Delta_{\mathbf{p}'}(\phi)$ is the standard deviation of the momentum distribution, where the points \mathbf{x}' and \mathbf{p}' are the corresponding expectation values. Moreover, using the fact $\mathbf{p} = \hbar\mathbf{k}$, we note that

$$(2.10) \quad \Delta_{\mathbf{p}'}(\phi) = \hbar\Delta_{\mathbf{p}'/\hbar}(\hat{\psi}) = \hbar\Delta_{\mathbf{k}'}(\hat{\psi}).$$

Now we can introduce the mathematical Heisenberg's uncertainty principle, see Ref. [6]:

If f and every component of ∇f belong to $L^2(\mathbb{R}^3; \mathbb{C})$ and $\int_{\mathbb{R}^3} |f(\mathbf{x})|^2 d\mathbf{x} = 1$, then

$$(2.11) \quad \Delta_{\mathbf{a}}(f)\Delta_{\mathbf{b}}(f) \geq \frac{3}{2}, \text{ for } \mathbf{a}, \mathbf{b} \in \mathbb{R}^3.$$

We will find out below that the above condition $\nabla\psi \in L^2(\mathbb{R}^3)$ implies that the expectation value of momentum is finite. As follows, we have the well-known formulation of the Heisenberg's uncertainty principle that illustrates the central role of the strictly positive Planck's constant. In the classical sense, the uncertainty principle can be seen as a statistical result for a very large ensemble of identical particles with the identical wave function, that is, if we measure position of one half of the particles and momentum of the other half, as accurately as technically possible, then the product of standard deviations has the sharp and positive lower bound:

$$(2.12) \quad \Delta_{\mathbf{x}'}(\psi)\Delta_{\mathbf{p}'}(\phi) \geq \frac{3}{2}\hbar, \text{ for } \mathbf{x}', \mathbf{p}' \in \mathbb{R}^3.$$

⁴The physically most typical Gaussian distribution has the same value for mean, median and mode. Therefore, the standard deviation is a natural measure of uncertainty when we consider a statistical physical phenomena.

Both the position and momentum of a particle are described by statistical distributions and due to the Heisenberg's uncertainty principle there are no well-defined values for both position and momentum at the same moment of time. Hence, the position distribution can be localized around multiple spatially distant points, rendering a very non-classical description of the location of a mass particle that is always a single-valued observable quantity. A classical (strong) measurement of the position equals the complete information and hence the term $|\psi|^2$ will vanish outside the observed location. At the very same time, the momentum will appear everywhere in the momentum space. Therefore, if the directly or indirectly gained information changes the information content, then all distributions are updated in a non-local fashion in order to preserve the natural probability condition, i.e., the sum of all probabilities remains a unity (cf. *objective Bayesian inference* and also Ref. [7]). A related phenomenon that illustrates the indirectly available information is the wave mechanical *entangled state*, see Ref. [5].

The peculiar fact that there is no definite value for the position of a particle allows to prevent a hydrogen-like atom from collapsing, an inevitable and paradoxal consequence of electrostatic forces in classical physics as the electron spirals into the nucleus and the atomic system releases an infinite amount of energy. Here the uncertainty principle predicts that there is a form of a steady state equilibrium between the electron and the massive nucleus that is actively localizing the electron by interaction that carries energy; the information content associated with the state of motion causes an entropic pressure that counters the force due to the localizing interaction.

Having the probability distributions allows us to consider expectation values for observable physical quantities that are dependent on position or momentum. In particular, the position \mathbf{x} of a particle, the real valued potential energy $V(\mathbf{x})$ of a particle and the corresponding external force $\mathbf{F}(\mathbf{x})$ acting on a particle have the following expectation values:

$$(2.13) \quad \mathbb{E}_\psi(\mathbf{x}) = \int_{\mathbb{R}^3} \mathbf{x} |\psi(\mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^3} \mathbf{x} \psi(\mathbf{x}) \overline{\psi(\mathbf{x})} d\mathbf{x};$$

$$(2.14) \quad \mathbb{E}_\psi(V) = \int_{\mathbb{R}^3} V(\mathbf{x}) |\psi(\mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^3} V(\mathbf{x}) \psi(\mathbf{x}) \overline{\psi(\mathbf{x})} d\mathbf{x};$$

$$(2.15) \quad \mathbb{E}_\psi(\mathbf{F}) = \mathbb{E}_\psi(-\nabla V) = - \int_{\mathbb{R}^3} \nabla V(\mathbf{x}) \psi(\mathbf{x}) \overline{\psi(\mathbf{x})} d\mathbf{x}.$$

Furthermore, as it is well-known, the expectation values of the momentum \mathbf{p} and the kinetic energy T can be represented in the position space, when we apply the Fourier transform of derivative and Plancherel's theorem:

$$(2.16) \quad \mathbb{E}_\phi(\mathbf{p}) = \int_{\mathbb{R}^3} \mathbf{p} |\phi(\mathbf{p})|^2 d\mathbf{p} = \int_{\mathbb{R}^3} [\mathbf{p}\phi(\mathbf{p})] \overline{\phi(\mathbf{p})} d\mathbf{p} = \int_{\mathbb{R}^3} [-i\hbar\nabla\psi(\mathbf{x})] \overline{\psi(\mathbf{x})} d\mathbf{x} = \mathbb{E}_\psi(\mathbf{p});$$

$$(2.17) \quad \mathbb{E}_\phi(T) = \int_{\mathbb{R}^3} \frac{\mathbf{p}^2}{2m} |\phi(\mathbf{p})|^2 d\mathbf{p} = \frac{1}{2m} \int_{\mathbb{R}^3} [-\hbar^2\nabla^2\psi(\mathbf{x})] \overline{\psi(\mathbf{x})} d\mathbf{x} = \mathbb{E}_\psi(T).$$

We also have the expectation value of the total energy E that is the following:

$$(2.18) \quad \mathbb{E}_\psi(E) = \int_{\mathbb{R}^3} \left(-\frac{\hbar^2}{2m} \nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) \right) \overline{\psi(\mathbf{x})} d\mathbf{x}.$$

In classical mechanics the angular momentum for a point-like particle with respect to the origin is $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. Its component-wise representation is $L_1 = x_2p_3 - x_3p_2$, $L_2 = x_3p_1 - x_1p_3$ and $L_3 = x_1p_2 - x_2p_1$. For example, the expectation value of x_1p_2 can be found if we consider the inverse Fourier transform of $p_2\phi(\mathbf{p})$ and then multiply both sides by x_1 . Hence,

$$-i\hbar \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \right) \psi(\mathbf{x}) = \frac{1}{\sqrt{2\pi\hbar^3}} \int_{\mathbb{R}^3} (x_1p_2 - x_2p_1) \phi(\mathbf{p}) \exp(i\langle \mathbf{p}, \mathbf{x} \rangle / \hbar) d\mathbf{p}.$$

Via the Plancherel's theorem we have the following expression for the expectation value:

$$(2.19) \quad \mathbb{E}_\phi(\mathbf{L}) = \int_{\mathbb{R}^3} (\mathbf{x} \times \mathbf{p}) |\phi(\mathbf{p})|^2 d\mathbf{p} = \int_{\mathbb{R}^3} [(-i\hbar \mathbf{x} \times \nabla) \psi(\mathbf{x})] \overline{\psi(\mathbf{x})} d\mathbf{x} = \mathbb{E}_\psi(\mathbf{L}).$$

From the previous considerations we can assume that every observable physical quantity O corresponds to a linear self-adjoint operator \hat{O} of $L^2(\mathbb{R}^3; \mathbb{C})$. This ensures that $\mathbb{E}_\psi(O)$ is real and the eigenvectors of \hat{O} span an orthonormal basis of $L^2(\mathbb{R}^3; \mathbb{C})$, see Ref. [6]. In turn, this basis allows us to represent the expectation value as a linear weighted sum of the corresponding real eigenvalues. In particular, a classical measurement of O must realize only one of the possible independent states, i.e., one of the eigenvectors of \hat{O} , and the result is the corresponding eigenvalue, see Ref. [5]. We also require that a physical expectation value must be finite, thus we can infer from the Schwarz's inequality that the domain of an operator \hat{O} is a linear subspace of $L^2(\mathbb{R}^3; \mathbb{C})$ where the expectation value integral exists, i.e.,

$$(2.20) \quad \mathfrak{D}(\hat{O}) := \{\psi \in L^2(\mathbb{R}^3; \mathbb{C}) : \hat{O}\psi \in L^2(\mathbb{R}^3; \mathbb{C})\} \subset L^2(\mathbb{R}^3; \mathbb{C}).$$

For this reason we must assume that set $\mathfrak{D}(\hat{O})$ is dense in $L^2(\mathbb{R}^3; \mathbb{C})$ in order to consider all the physical systems. The analysis of previous classical operators can be found in Ref. [6].

Next we consider a situation where the probability distributions that express the state of motion of a particle change in time, i.e., the wave function is of the form $\psi(\mathbf{x}, t)$. We note that time is not an observable physical quantity that has a probability distribution such as variables \mathbf{x} and \mathbf{p} , but it is only a classically progressive parameter associated to any evolving system. We note also that we should avoid unnecessary constraints to keep the configuration space as free as possible and allow an entropy driven physical model to take its form. Now, in the case of time-evolution, it is not obvious how to formulate the corresponding *Liouville equation* of statistical classical mechanics for the wave mechanical phase space distribution, but we may certainly assume that the expectation values obey the *Newton's equations of motion*:

$$(2.21) \quad m \frac{d}{dt} \mathbb{E}_\psi(\mathbf{x}) = \mathbb{E}_\psi(\mathbf{p});$$

$$(2.22) \quad \frac{d}{dt} \mathbb{E}_\psi(\mathbf{p}) = \mathbb{E}_\psi(\mathbf{F}).$$

We note that the statistical model must exist all the time (cf. *Liouville's theorem*), i.e., since Eqs. (2.21) and (2.22) belong to the position space, via the crucial Eq. (2.16), we have

$$(2.23) \quad \frac{d}{dt} \int_{\mathbb{R}^3} |\psi(\mathbf{x}, t)|^2 d\mathbf{x} = 0,$$

and the value of the integral is a unity, independently of time $t \geq 0$. When we try to solve ψ from the previous integral equations the task is not trivial, but, for historical reasons, we have a guideline how to proceed, see the *Ehrenfest's theorem* in Refs. [5; 8]. Since we consider a physical particle that is described by its wave function ψ , it is evident that the expectation values of position, momentum, kinetic energy, potential energy and force exist, i.e., all the following terms ψ , $\mathbf{x}\psi$, $\nabla\psi$, $\nabla^2\psi$, V , $V\psi$ and $\nabla V\psi$ are continuous and vanish as $|\mathbf{x}| \rightarrow \infty$. Also, the statistical condition (2.23) demands a suitable expression for the continuous and bounded⁵ time derivative $\partial_t\psi$. With these considerations, let us begin with the expectation value of force and use partial integration when trying to solve the time derivative:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}_\psi(\mathbf{p}) &= \int_{\mathbb{R}^3} \frac{\partial}{\partial t} ((-i\hbar \nabla \psi) \bar{\psi}) d\mathbf{x} = \int_{\mathbb{R}^3} \left(-i\hbar \frac{\partial}{\partial t} \bar{\psi} \nabla \psi - i\hbar \bar{\psi} \frac{\partial}{\partial t} (\nabla \psi) \right) d\mathbf{x} \\ &= \int_{\mathbb{R}^3} \left(-i\hbar \frac{\partial}{\partial t} \bar{\psi} \nabla \psi + i\hbar \frac{\partial}{\partial t} \psi \nabla \bar{\psi} \right) d\mathbf{x} = \int_{\mathbb{R}^3} 2\text{Re} \left[i\hbar \frac{\partial}{\partial t} \psi \nabla \bar{\psi} \right] d\mathbf{x}. \end{aligned}$$

⁵A physical process that updates the information content is not instant by its nature.

On the other hand, in aiming to eliminate the variable term $\nabla\bar{\psi}$, we have

$$\begin{aligned}\mathbb{E}_\psi(\mathbf{F}) &= \mathbb{E}_\psi(-\nabla V) = \int_{\mathbb{R}^3} -\nabla V |\psi|^2 d\mathbf{x} = \int_{\mathbb{R}^3} V \nabla |\psi|^2 d\mathbf{x} \\ &= \int_{\mathbb{R}^3} (V\bar{\psi}\nabla\psi + V\psi\nabla\bar{\psi}) d\mathbf{x} = \int_{\mathbb{R}^3} 2\text{Re} [V\psi\nabla\bar{\psi}] d\mathbf{x}.\end{aligned}$$

Thus, the equality holds in Eq. (2.22) if $\partial_t\psi = -iV\psi/\hbar$ that represents the potential energy.

Next we consider the expectation value of the first component of the momentum vector:

$$\begin{aligned}m \frac{d}{dt} [\mathbb{E}_\psi(\mathbf{x})]_1 &= \int_{\mathbb{R}^3} m \frac{\partial}{\partial t} (x_1 |\psi|^2) d\mathbf{x} = \int_{\mathbb{R}^3} \left(mx_1\psi \frac{\partial}{\partial t} \bar{\psi} + mx_1\bar{\psi} \frac{\partial}{\partial t} \psi \right) d\mathbf{x} \\ &= \int_{\mathbb{R}^3} 2\text{Re} \left[mx_1\bar{\psi} \frac{\partial}{\partial t} \psi \right] d\mathbf{x}.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}[\mathbb{E}_\psi(\mathbf{p})]_1 &= \int_{\mathbb{R}^3} \left(-i\hbar \frac{\partial\psi}{\partial x_1} \right) \bar{\psi} d\mathbf{x} \\ &= -\frac{i\hbar}{2} \int_{\mathbb{R}^3} \left(\bar{\psi} \frac{\partial\psi}{\partial x_1} - \psi \frac{\partial\bar{\psi}}{\partial x_1} \right) d\mathbf{x} \\ &= -\frac{i\hbar}{2} \int_{\mathbb{R}^3} \left(\bar{\psi} \frac{\partial\psi}{\partial x_1} - \psi \frac{\partial\bar{\psi}}{\partial x_1} \right) d\mathbf{x} + \frac{i\hbar}{2} \int_{\mathbb{R}^3} \nabla \cdot (x_1\bar{\psi}\nabla\psi - x_1\psi\nabla\bar{\psi}) d\mathbf{x} \\ &= -\frac{i\hbar}{2} \int_{\mathbb{R}^3} \left(\bar{\psi} \frac{\partial\psi}{\partial x_1} - \psi \frac{\partial\bar{\psi}}{\partial x_1} \right) d\mathbf{x} \\ &\quad + \frac{i\hbar}{2} \int_{\mathbb{R}^3} (x_1\nabla\bar{\psi} \cdot \nabla\psi + \bar{\psi}\mathbf{e}_1 \cdot \nabla\psi + x_1\bar{\psi}\nabla^2\psi) d\mathbf{x} \\ &\quad - \frac{i\hbar}{2} \int_{\mathbb{R}^3} (x_1\nabla\psi \cdot \nabla\bar{\psi} + \psi\mathbf{e}_1 \cdot \nabla\bar{\psi} + x_1\psi\nabla^2\bar{\psi}) d\mathbf{x} \\ &= \int_{\mathbb{R}^3} \left(\frac{i\hbar}{2} x_1\bar{\psi}\nabla^2\psi - \frac{i\hbar}{2} x_1\psi\nabla^2\bar{\psi} \right) d\mathbf{x} \\ &= \int_{\mathbb{R}^3} 2\text{Re} \left[\frac{i\hbar}{2} x_1\bar{\psi}\nabla^2\psi \right] d\mathbf{x}.\end{aligned}$$

Thus, the equality in Eq. (2.21) holds for every component if $\partial_t\psi = i\hbar\nabla^2\psi/2m$ that represents the kinetic energy. The additional integral over divergence is zero due to the Gauss' theorem since the integrand vanishes outside some large sphere in \mathbb{R}^3 .

Furthermore, if the time-evolution of a state depends only on a spatial position, then we can infer that the effect of momentum is absent. Indeed, if $\partial_t\psi = -iV\psi/\hbar$, then

$$m \frac{d}{dt} \mathbb{E}_\psi(\mathbf{x}) = \int_{\mathbb{R}^3} \left(m\mathbf{x}\psi \frac{\partial}{\partial t} \bar{\psi} + m\mathbf{x}\bar{\psi} \frac{\partial}{\partial t} \psi \right) d\mathbf{x} = \int_{\mathbb{R}^3} \left(m\mathbf{x}\psi \frac{i}{\hbar} V\bar{\psi} - m\mathbf{x}\bar{\psi} \frac{i}{\hbar} V\psi \right) d\mathbf{x} = 0.$$

On the other hand, if the time-evolution of a state depends only on a spatial translation, then we can infer that the effect of external force is absent. Thus, if $\partial_t\psi = i\hbar\nabla^2\psi/2m$, then we can consider the expectation value of the first component of the force vector:

$$\begin{aligned}\frac{d}{dt} [\mathbb{E}_\psi(\mathbf{p})]_1 &= \int_{\mathbb{R}^3} \left(-i\hbar \frac{\partial}{\partial t} \bar{\psi} \frac{\partial\psi}{\partial x_1} + i\hbar \frac{\partial}{\partial t} \psi \frac{\partial\bar{\psi}}{\partial x_1} \right) d\mathbf{x} = \int_{\mathbb{R}^3} \left(-\frac{\hbar^2}{2m} \nabla^2\bar{\psi} \frac{\partial\psi}{\partial x_1} - \frac{\hbar^2}{2m} \nabla^2\psi \frac{\partial\bar{\psi}}{\partial x_1} \right) d\mathbf{x} \\ &= -\frac{\hbar^2}{2m} \int_{\mathbb{R}^3} \nabla \cdot \left(\nabla\bar{\psi} \frac{\partial\psi}{\partial x_1} + \nabla\psi \frac{\partial\bar{\psi}}{\partial x_1} \right) d\mathbf{x} = 0.\end{aligned}$$

The previous result is the same for the other components as well. Therefore, due to linearity, a choice $\partial_t \psi = i\hbar \nabla^2 \psi / 2m - iV\psi/\hbar$ satisfies the both Eqs. (2.21) and (2.22). Finally, with the previous expression for the time derivative, Eq. (2.23) is also satisfied:

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^3} |\psi|^2 d\mathbf{x} &= \int_{\mathbb{R}^3} \left(\psi \left(\frac{-i\hbar}{2m} \nabla^2 \bar{\psi} + \frac{i}{\hbar} V \bar{\psi} \right) + \bar{\psi} \left(\frac{i\hbar}{2m} \nabla^2 \psi - \frac{i}{\hbar} V \psi \right) \right) d\mathbf{x} \\ &= \frac{i\hbar}{2m} \int_{\mathbb{R}^3} \nabla \cdot (\bar{\psi} \nabla \psi - \psi \nabla \bar{\psi}) d\mathbf{x} = 0. \end{aligned}$$

To sum up, we have arrived to the following fundamental result:

$$(2.24) \quad i\hbar \frac{\partial \psi}{\partial t}(\mathbf{x}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{x}, t) + V(\mathbf{x}, t) \psi(\mathbf{x}, t), \text{ for } (\mathbf{x}, t) \in \mathbb{R}^3 \times [0, \infty[.$$

This is the *Schrödinger equation* that describes the time-evolution of ψ when a mass particle interacts with an external force field. The Heisenberg's uncertainty principle manifests itself as follows: a binding force localizes the particle (the compression of ψ) but the momentum and the kinetic energy will become more uncertain (the spread of ϕ). Therefore, ψ is in the sense of maximum entropy the most probable state that satisfies the conservation of energy, i.e.,

$$(2.25) \quad \mathbb{E}_\psi(E) - \mathbb{E}_\psi(T) - \mathbb{E}_\psi(V) = \int_{\mathbb{R}^3} \left[i\hbar \frac{\partial \psi}{\partial t} + \frac{\hbar^2}{2m} \nabla^2 \psi - V\psi \right] \bar{\psi} d\mathbf{x} = 0,$$

where the time derivative is associated with the total energy E . By definition, the epistemic state ψ is a prior to any further interaction that may change the information content of the system. Since the equation is of the first order in time, the initial condition $\psi(\mathbf{x}, t_0)$ determines the time-evolution of the probability distributions of observables. Every solution $\psi(\mathbf{x}, t)$ evolves continuously and deterministically but a change in the information content can change the wave function $\psi(\mathbf{x}, t)$ in a seemingly non-continuous fashion. In particular, a classical measurement of the position equals the complete information and the associated physical process causes an extremely rapid concentration of the wave function around the observed location. Consequently, a new initial state $\psi(\mathbf{x}, t_0)$ emerges that again evolves according to the Schrödinger equation. The equation is also linear, building a far-reaching connection with the Huygens' principle that describes the wave propagation and allows to explain interference-like observations. Lastly, if the position distribution is time-independent, i.e., $|\psi(\mathbf{x}, t)|^2 = |\psi(\mathbf{x})|^2$, then we can conclude that there is no change in entropy of the joint state distribution, see the related *stationary state* in Ref. [5] that allows to explain – to some extent – the atomic spectra and stability.

In wave mechanics the coupled position and momentum distributions are spread out around the average values, like a statistical ensemble of particles, and an entropic force drives the joint distribution towards the maximum randomness. Incidentally, the compression of ψ creates a form of an internal kinetic energy, which increase equals to the work done by the environment in gaining information via physical interactions. Conversely, the spread of ψ is an entropy driven process that erases information by dissipating internal kinetic energy into the environment. Along these lines, we have also noticed that the classical trajectory $(\mathbf{x}(t), \mathbf{p}(t))$ is replaced by the wave function $\psi(\mathbf{x}, t)$ that contains all the position and momentum dependent observable information. To this end, we have the expectation value of a generic observable $O(\mathbf{x}, \mathbf{p}, t)$, i.e.,

$$(2.26) \quad \mathbb{E}_\psi(O) = \int_{\mathbb{R}^3} \overline{\psi(\mathbf{x}, t)} \hat{O} \psi(\mathbf{x}, t) d\mathbf{x},$$

and the related operators that correspond physical quantities in classical mechanics:

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{x}, \quad \hat{\mathbf{p}} = -i\hbar \nabla, \quad \hat{\mathbf{L}} = -i\hbar \mathbf{x} \times \nabla \quad \text{and} \quad \hat{\mathbf{F}} = -\nabla V(\mathbf{x}, t); \\ \hat{V} &= V(\mathbf{x}, t), \quad \hat{T} = -\frac{\hbar^2}{2m} \nabla^2 \quad \text{and} \quad \hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}, t). \end{aligned}$$

3. CONCLUSION

Entropy can be understood as a measure of uncertainty, disorder or missing information content of a system. E. T. Jaynes introduced the information theoretical maximum entropy principle to have a least biased estimate possible on the given information in the framework of statistical physics [4], i.e., the most probable statistical distribution of microstates is the least biased estimate that is consistent with all the macroscopic constraints that determine the particular physical model. Since the state of motion is an inherently probabilistic concept, the maximized entropy corresponds to the most probable statistical distribution of microstates before a measurement. Here we have obtained the non-relativistic wave mechanical description of a single, interacting mass particle with no internal degrees of freedom by using the maximum entropy principle as a basis of modelling together with a set of relevant physical constraints that arise from the non-classical uncertainty principle and well-established classical theory.

There are also other probability estimation based approaches in the literature, but they differ from our quite straightforward presentation that is an amalgamation of many known results. It is also worth mentioning that in the famous article, Ref. [3], W. Heisenberg addressed his desire to derive the rules of wave mechanics directly from the physical foundations, i.e., from the uncertainty principle. Here we have demonstrated that the Heisenberg's uncertainty principle, in the context of information theory of partial knowledge, is the decisive ingredient that allows to derive the dynamical law of motion in the wave mechanical theory. However, the intrinsically random and non-local character of the theory remains an unexplained physical phenomenon.

ACKNOWLEDGEMENTS

Here I would like to thank professor Dr. Esa Räsänen and M.Sc. Joonas Keski-Rahkonen for reading the present work and encouraging its publication.

REFERENCES

- [1] E. Schrödinger, "Quantisierung als Eigenwertproblem". *Annalen der Physik* 384, (4), 361-376, (1926); *Annalen der Physik* 384, (6), 489-527, (1926); *Annalen der Physik* 385, (13), 437-490, (1926); *Annalen der Physik* 386, (18), 109-139, (1926).
- [2] M.J.W. Hall, M. Reginatto, "Schrödinger equation from an exact uncertainty principle". *Journal of Physics A: Mathematical and General* 35, (14), 3289-3303, (2002).
- [3] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". *Zeitschrift für Physik* 43, (3-4), 172-198, (1927).
- [4] E.T. Jaynes, "Information Theory and Statistical Mechanics". *Physical Review* 106, (4), 620-630, (1957); *Physical Review* 108, (2), 171-190, (1957).
- [5] E. Merzbacher, *Quantum Mechanics*. Wiley, 3 edition, (1997).
- [6] G. Teschl, *Mathematical Methods in Quantum Mechanics: With Applications to Schrödinger Operators*. American Mathematical Society, 2 edition, (2014).
- [7] E.T. Jaynes, "Probability in quantum theory" in *Complexity, Entropy and the Physics of Information* (ed. W.H. Zurek). Addison-Wesley, Redwood City, CA, (1990).
- [8] P. Ehrenfest, "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik". *Zeitschrift für Physik* 45, (7-8), 455-457, (1927).