A maximum entropy approach to wave mechanics

Juho Leppäkangas[1]

30.8.2020

ABSTRACT.

We employ the maximum entropy principle, in the context of statistical inference by impersonal physical interactions, together with the experimental position-momentum uncertainty phenomenon to construct 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.

[1]Email: 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 the experimental uncertainty phenomenon, 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 representation.

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 [5-6]. 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 phenomenon 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 particle 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.[3] 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.

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, and 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 [3; 5].

The time-independent reciprocality between position and momentum distributions suggests that there exists an invertible transformation that scales two weight functions in a reciprocal manner, and the characteristical wave phenomenon motivates to seek a linear transformation. In the spirit of the maximum entropy principle [4], we begin by forming unbiased estimations of the both distributions that model the time-independent random process and reciprocality. Our preliminary model is a probabilistic field, a weighted scalar sum in a respective space:

---

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

[3] Perhaps this information is encoded and stored within the other physical subsystems of the Universe.

$$\psi(\mathbf{x}) = \sum_{\mathbf{p}\in\mathbb{R}^3} \phi(\mathbf{p}) v_{\mathbf{x}}(\mathbf{p}) \quad \text{and} \quad \sum_{\mathbf{x}\in\mathbb{R}^3} \|\psi(\mathbf{x})\|^2 = 1, \tag{2.1}$$

$$\phi(\mathbf{p}) = \sum_{\mathbf{x}\in\mathbb{R}^3} \psi(\mathbf{x}) u_{\mathbf{p}}(\mathbf{x}) \quad \text{and} \quad \sum_{\mathbf{p}\in\mathbb{R}^3} \|\phi(\mathbf{p})\|^2 = 1, \tag{2.2}$$

where the functions $\|\psi\|^2$ and $\|\phi\|^2$ give respectively the position and momentum distributions. Here the functions $v_{\mathbf{x}}(\cdot)$ and $u_{\mathbf{p}}(\cdot)$ are suitable scale factors that determine the reciprocality, and the common vector space norm $\|\cdot\|$ is assumed to be induced by a suitable inner product. Thus, the ill-defined functions ψ and ϕ are reciprocal to each other: if $\psi(\mathbf{x})$ vanishes outside of a narrow interval $\Delta\mathbf{x}$, then $\phi(\mathbf{p})$ must be non-zero on a large interval $\Delta\mathbf{p}$ and vice versa. Also, the observed intensity $\|\phi(\mathbf{p})\|^2$ should not change under a constant shift $\psi(\mathbf{x}) \to \psi(\mathbf{x}+\mathbf{x}_0)$.

Now we assume that there is an invertible linear transformation F between the probabilistic "weight vectors" $\{\psi(\mathbf{x})\}$ and $\{\phi(\mathbf{p})\}$ such that $F\{\psi(\mathbf{x})\} = \{\phi(\mathbf{p})\}$, and if we consider the norms of probabilistic unit vectors, then we can also assume that F is a unitary transformation. Moreover, the both of ψ and ϕ can be interpreted as an inner product of a probabilistic column vector and one of the orthonormal row vectors. In order to conclude, we assume that the both ψ and ϕ are continuous functions and F is a unitary transformation in a suitable function space, and due to the wave-like characteristics of F, the orthonormal basis vectors $v_{\mathbf{x}}(\cdot)$ and $u_{\mathbf{p}}(\cdot)$ possess periodic properties. 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*:

$$\psi(\mathbf{x}) = \int_{\mathbb{R}^3} \phi(\mathbf{p}) v_{\mathbf{x}}(\mathbf{p}) d\mathbf{p} \quad \text{and} \quad \phi(\mathbf{p}) = \int_{\mathbb{R}^3} \psi(\mathbf{x}) u_{\mathbf{p}}(\mathbf{x}) d\mathbf{x}. \tag{2.3}$$

From harmonic analysis, and also from the theory of optical diffraction [6], 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 spectrums and intensities of the two random variables.[4] 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 all vectors $\exp(i\mathbf{k}\cdot\mathbf{x})/\sqrt{2\pi}^3$ or vectors $\exp(-i\mathbf{k}\cdot\mathbf{x})/\sqrt{2\pi}^3$. Since both the position and momentum of a physical particle are finite, the both functions $\psi(\mathbf{x})$ and $\phi(\mathbf{p})$ must vanish as $|\mathbf{x}|, |\mathbf{p}| \to \infty$, i.e., low entropy microstates are negligible, and we may assume 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. [7]:

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

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:

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

where the components of $\mathbf{q}$ have the dimension of [1/momentum]. Here we employ the L^2-norm to probabilistic weights and hence the spectral densities $|\check{\phi}|^2$ and $|\hat{\psi}|^2$ correspond respectively to position and momentum distributions. Due to the Plancherel's theorem, we have

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

[4]The 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., $|\psi_1 + i\psi_2|^2 = \psi_1^2 + \psi_2^2$. From an algebraic point of view, the Fourier transform behaves like a change of basis matrix, and the unitarity allows to connect two probabilistic sums.

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 [length $\times$ momentum] = [time $\times$ 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:

$$\psi(\mathbf{x}) = \frac{1}{\sqrt{2\pi\hbar}^3} \int_{\mathbb{R}^3} \phi(\mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{x}/\hbar) d\mathbf{p}; \tag{2.6}$$

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

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 probabilistic 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 model for the uncertainty:[5]

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

$$\triangle_{\mathbf{a}}(f) := \sqrt{\int_{\mathbb{R}^3} |\mathbf{x} - \mathbf{a}|^2 \, |f(\mathbf{x})|^2 \, d\mathbf{x}}. \tag{2.8}$$

It is clear that the deviation $\triangle_{\mathbf{a}}(f)$ 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, $\triangle_{\mathbf{x}'}(\psi)$ is the standard deviation of the position distribution and $\triangle_{\mathbf{p}'}(\phi)$ is the standard deviation of the momentum distribution, where the points $\mathbf{x}'$ and $\mathbf{p}'$ are the corresponding expectation values. Moreover, after a change of variable, we note that

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

Now we can introduce the mathematical Heisenberg's uncertainty principle, see Ref. [7]:

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

$$\triangle_{\mathbf{a}}(f) \triangle_{\mathbf{b}}(\hat{f}) \geq \frac{3}{2}, \quad \text{for } \mathbf{a}, \mathbf{b} \in \mathbb{R}^3. \tag{2.9}$$

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 positive – and sharp – lower bound:

$$\triangle_{\mathbf{x}'}(\psi) \triangle_{\mathbf{p}'}(\phi) \geq \frac{3}{2}\hbar, \quad \text{for } \mathbf{x}', \mathbf{p}' \in \mathbb{R}^3. \tag{2.10}$$

---

[5] The 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 statistical physical phenomena. On the other hand, let us note that in the framework of information theory, both the Fisher and Shannon information measures can be used to produce meaningful uncertainty relations due to ψ and ϕ, see Ref. [14].

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 variables simultaneously. Also, 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 (ideal or 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 appears everywhere in the momentum space, i.e., it is completely indefinite. Moreover, if the directly or indirectly exerted interaction changes the information content, then every distribution is 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 Refs. [8-9]. A phenomenon that illustrates the indirectly available part of an information content is the wave mechanical *entangled state* [5; 9].

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. The uncertainty principle allows a form of steady state equilibrium between the electron and massive nucleus that is actively localizing the electron by interaction that carries energy; the information content associated with a state of motion causes an entropic pressure that counters the electric force that, in turn, contributes to the information content.

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

$$\mathbb{E}_\psi(\mathbf{x}) = \int_{\mathbb{R}^3} \mathbf{x}\,|\psi(\mathbf{x})|^2\, d\mathbf{x} = \int_{\mathbb{R}^3} \mathbf{x}\psi(\mathbf{x})\overline{\psi(\mathbf{x})}d\mathbf{x}; \tag{2.11}$$

$$\mathbb{E}_\psi(V) = \int_{\mathbb{R}^3} V(\mathbf{x})\,|\psi(\mathbf{x})|^2\, d\mathbf{x} = \int_{\mathbb{R}^3} V(\mathbf{x})\psi(\mathbf{x})\overline{\psi(\mathbf{x})}d\mathbf{x}; \tag{2.12}$$

$$\mathbb{E}_\psi(\mathbf{F}) = \mathbb{E}_\psi(-\nabla V) = -\int_{\mathbb{R}^3} \nabla V(\mathbf{x})\psi(\mathbf{x})\overline{\psi(\mathbf{x})}d\mathbf{x}. \tag{2.13}$$

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

$$\mathbb{E}_\phi(\mathbf{p}) = \int_{\mathbb{R}^3} \mathbf{p}\,|\phi(\mathbf{p})|^2\, d\mathbf{p} = \int_{\mathbb{R}^3} [\mathbf{p}\phi(\mathbf{p})]\,\overline{\phi(\mathbf{p})}d\mathbf{p} = \int_{\mathbb{R}^3} [-i\hbar\nabla\psi(\mathbf{x})]\,\overline{\psi(\mathbf{x})}d\mathbf{x} = \mathbb{E}_\psi(\mathbf{p}); \tag{2.14}$$

$$\mathbb{E}_\phi(T) = \int_{\mathbb{R}^3} \frac{\mathbf{p}^2}{2m}\,|\phi(\mathbf{p})|^2\, d\mathbf{p} = \frac{1}{2m}\int_{\mathbb{R}^3} \left[-\hbar^2\nabla^2\psi(\mathbf{x})\right]\overline{\psi(\mathbf{x})}d\mathbf{x} = \mathbb{E}_\psi(T). \tag{2.15}$$

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

$$\mathbb{E}_\psi(E) = \int_{\mathbb{R}^3} \left(-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x})\right)\overline{\psi(\mathbf{x})}d\mathbf{x}. \tag{2.16}$$

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 componentwise 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,

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

Via the Plancherel's theorem, we have the usual expression:

$$\mathbb{E}_\phi(\mathbf{L}) = \int_{\mathbb{R}^3} (\mathbf{x} \times \mathbf{p}) \, |\phi(\mathbf{p})|^2 \, d\mathbf{p} = \int_{\mathbb{R}^3} \left[(-i\hbar \mathbf{x} \times \nabla) \, \psi(\mathbf{x}) \right] \overline{\psi(\mathbf{x})} d\mathbf{x} = \mathbb{E}_\psi(\mathbf{L}). \tag{2.17}$$

We may assume from the previous considerations 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. [7]. In turn, this basis allows to represent the expectation value as a linear weighted sum of the corresponding real eigenvalues. In particular, a classical measurement of O can realize only one of the possible independent and classically exclusive states, i.e., one of the eigenvectors of $\hat{O}$, and the result is the corresponding eigenvalue [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 (componentwise) a subspace of $L^2(\mathbb{R}^3;\mathbb{C})$ where the expectation value integral exists, i.e.,

$$\mathfrak{D}(\hat{O}) := \{\psi \in L^2(\mathbb{R}^3;\mathbb{C}) \, : \, \hat{O}\psi \in L^2(\mathbb{R}^3;\mathbb{C})\} \subset L^2(\mathbb{R}^3;\mathbb{C}). \tag{2.18}$$

For this reason, we assume that set $\mathfrak{D}(\hat{O})$ is dense in $L^2(\mathbb{R}^3;\mathbb{C})$ in order to consider the set of all physical systems. The analysis of previous classical operators can be found in Ref. [7].

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 progressive parameter assosiated to any (entropically) evolving system. We note also that we should avoid unnecessary constraints to keep the configuration space as free as possible and thus 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 classical statistical mechanics for a wave mechanical phase space distribution, but we may certainly assume that expectation values obey the Newton's equations of motion:

$$m\frac{d}{dt}\mathbb{E}_\psi(\mathbf{x}) = \mathbb{E}_\psi(\mathbf{p}); \tag{2.19}$$

$$\frac{d}{dt}\mathbb{E}_\psi(\mathbf{p}) = \mathbb{E}_\psi(\mathbf{F}). \tag{2.20}$$

The correct model must provide not only the conservation of energy but also of probability. Thus, since Eqs. (2.19) and (2.20) are in the position space, via the crucial Eq. (2.14), we have

$$\frac{d}{dt}\int_{\mathbb{R}^3} |\psi(\mathbf{x},t)|^2 \, d\mathbf{x} = 0, \tag{2.21}$$

and the value of integral is a unity, independently of time $t \geq 0$. When we attempt to solve ψ from the previous integral equations the task is not trivial. Fortunately, we have a cue how to proceed, see the *Ehrenfest's theorem* in Refs. [5; 10]. See also Ref. [11]. Since we consider a physical particle, that is described by its wave function, 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 $|\mathbf{x}| \to \infty$. Also, the statistical condition in Eq. (2.21) demands a suitable expression for the continuous and bounded[6] time derivative $\partial_t\psi$. With these regularity conditions, let us now begin with the expectation value of force and use partial integration when trying to solve the time derivative:

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

[6] A physical process that updates the information content is not instant by its nature.

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

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

Thus, the equality holds in Eq. (2.20) if $\partial_t\psi = -iV\psi/\hbar$ that represents the potential energy. The previous choice for $\partial_t\psi$ is not unique, but here we follow equations with physical meaning. Next we consider the expectation value of the first component of momentum vector:

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

On the other hand, we have

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

Thus, the equality in Eq. (2.19) 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 an energy 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

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

On the other hand, if the time-evolution of an energy 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 force vector:

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

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. (2.19) and (2.20). Moreover, with the previous choice, the conservation of probability, Eq. (2.21), is also satisfied:

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

To sum up, we have arrived to the fundamental result for regular wave functions:

$$i\hbar\frac{\partial\psi}{\partial t}(\mathbf{x},t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t) + V(\mathbf{x},t)\psi(\mathbf{x},t), \text{ for } (\mathbf{x},t)\in\mathbb{R}^3\times[0,\infty[. \tag{2.22}$$

This is the *Schrödinger equation* that describes the system's time-evolution 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 then the momentum becomes more uncertain (the spread of ϕ) and kinetic energy increases. Thus, ψ is in the sense of maximum entropy the most probable state that satisfies the conservation of energy, i.e.,

$$\mathbb{E}_\psi(E) - \mathbb{E}_\psi(T) - \mathbb{E}_\psi(V) = \int_{\mathbb{R}^3}\left[i\hbar\frac{\partial\psi}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\psi - V\psi\right]\overline{\psi}d\mathbf{x} = 0, \tag{2.23}$$

where the time derivative is associated with the total energy E. By definition, the epistemic state ψ is a prior to any extra 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 probability distributions of all observables. Every solution $\psi(\mathbf{x},t)$ evolves continuously and deterministically but a change in the information content can change $\psi(\mathbf{x},t)$ in a seemingly non-continuous fashion. In particular, a classical measurement of the position equals the complete information and the associated physical process of updating causes an extremely rapid consentration 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 entropic change in the joint state distribution, see the related *stationary state* in Ref. [5] that allows to explain – to some extent – the atomic spectra and stability.

The energy of a physical system tends to disperse spontaneously, i.e., the macrostate becomes more homogeneous. In wave mechanics the coupled position and momentum distributions are spread out around the average values and an entropic force drives the joint distribution towards the maximum randomness within all the macroscopic constraints. Incidentally, the compression of ψ creates a form of an internal kinetic energy, which increase equals to the work done by the enviroment in gaining observable information via physical interactions. Conversely, the spread of ψ is an entropy driven process that erases information by dissipating kinetic energy into the enviroment, cf. Ref. [12]. In this physical model 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.,

$$\mathbb{E}_\psi(O) = \int_{\mathbb{R}^3}\overline{\psi(\mathbf{x},t)}\hat{O}\psi(\mathbf{x},t)d\mathbf{x}, \tag{2.24}$$

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

$$\begin{gathered}\hat{\mathbf{x}} = \mathbf{x},\quad \hat{\mathbf{p}} = -i\hbar\nabla,\quad \hat{\mathbf{L}} = -\,i\hbar\mathbf{x}\times\nabla \ \text{ and } \ \hat{\mathbf{F}} = -\nabla V(\mathbf{x},t); \\ \hat{V} = V(\mathbf{x},t),\quad \hat{T} = -\frac{\hbar^2}{2m}\nabla^2 \ \text{ and } \ \hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x},t).\end{gathered} \tag{2.25}$$

3. Conclusions

Entropy can be understood as a measure of disorder, uncertainty or missing information content of a system. E. T. Jaynes introduced the information theoretical maximum entropy principle to have the least biased estimate possible on a 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: the Heisenberg's uncertainty principle and its representation through de Broglie's relation and Heisenberg–Weyl inequality, Copenhagen interpretation with Bayesianism, Dirac–von Neumann axioms, Born's conditions and Ehrenfest's theorem. The analogous works are built upon models that contain statistical constraints for dynamical variables, see e.g., Refs. [2; 11; 13-21], that are related to the uncertainty principle. It may also be worth mentioning that in the famous article, Ref. [3], Heisenberg addressed his desire to derive the rules of wave mechanics (rather, matrix mechanics) directly from the physical foundations, e.g., 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 model. It should also be stressed that all the predictions of wave mechanics are complete in the strict sense that there is a fundamental limit what can be known about the system that is to be observed. Let us conclude that the intrinsically random and non-local character of the theory remains an unexplained phenomenon that seems to imply that all the individual events of Nature are subject to transcendental information processing.

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 online 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 Kinematic 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, 3rd edition, (1997).

[6] E. Hecht, Optics. Pearson, 5th edition, (2016).

[7] G. Teschl, Mathematical Methods in Quantum Mechanics: With Applications to Schrödinger Operators. American Mathematical Society, 2nd edition, (2014).

[8] E.T. Jaynes, “Probability in quantum theory” in Complexity, Entropy and the Physics of Information (ed. W.H. Zurek), 381-404. Addison-Wesley, Redwood City, CA, (1990).

[9] Č. Brukner, A. Zeilinger, “Information and Fundamental Elements of the Structure of Quantum Theory” in Time, Quantum and Information (ed. L. Castell, O. Ischebeck), 323-354. Springer, Berlin, Heidelberg, (2003).

[10] P. Ehrenfest, “Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik”. Zeitschrift für Physik 45, (7-8), 455-457, (1927).

[11] D.I. Bondar, R. Cabrera, R.R. Lompay, M.Y. Ivanov, H.A. Rabitz, “Operational Dynamic Modeling Transcending Quantum and Classical Mechanics”. Physical Review Letters 109, (19), 190403, (2012).

[12] R. Landauer, “Irreversibility and Heat Generation in the Computing Process”. IBM Journal of Research and Development 5, (3), 183-191, (1961).

[13] B.R. Frieden, “Fisher information as the basis for the Schrödinger wave equation”. American Journal of Physics 57, (11), 1004-1008, (1989).

[14] B.R. Frieden, Physics from Fisher information. Cambridge University Press, (1998).

[15] M. Reginatto, “Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information”. Physical Review A 58, (3), 1775-1778, (1998).

[16] F.H. Fröhner, “Missing Link Between Probability Theory and Quantum Mechanics: the Riesz-Fejer Theorem”. Zeitschrift für Naturforschung A 53, (8), 637-654, (1998).

[17] G. Grössing, “The vacuum fluctuation theorem: Exact Schrödinger equation via nonequilibrium thermodynamics”. Physics Letters A 372, (25), 4556-4563, (2008).

[18] U. Klein, "The statistical origins of quantum mechanics". Physics Research International 2010, 808424, (2010).

[19] A. Caticha, “Entropic dynamics, time and quantum theory”. Journal of Physics A: Mathematical and Theoretical 44, (22), 225303, (2011).

[20] S.D. Bartlett, T. Rudolph, R.W. Spekkens, “Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction”. Physical Review A 86, (1), 012103, (2012).

[21] A. Budiyono, D. Rohrlich, “Quantum mechanics as classical statistical mechanics with an ontic extension and an epistemic restriction”. Nature Communications 8, 1306, (2017).