Review

Entropic Dynamics

Ariel Caticha

Physics Department, University at Albany-SUNY, Albany, NY 12222, USA;
E-Mail: ariel@albany.edu; Tel.: +1-518-442-4592

Abstract: Entropic Dynamics is a framework in which dynamical laws are derived as an application of entropic methods of inference. No underlying action principle is postulated. Instead, the dynamics is driven by entropy subject to the constraints appropriate to the problem at hand. In this paper we review three examples of entropic dynamics. First we tackle the simpler case of a standard diffusion process which allows us to address the central issue of the nature of time. Then we show that imposing the additional constraint that the dynamics be non-dissipative leads to Hamiltonian dynamics. Finally, considerations from information geometry naturally lead to the type of Hamiltonian that describes quantum theory.

1. Introduction

The laws of physics, and in particular the laws of dynamics, have traditionally been seen as laws of nature. It is usually believed that such laws are *discovered* and that they are useful because they *reflect* reality. The reflection, imperfect though it may be, represents a very direct relation between physics and nature. Here we explore an alternative view in which the relation is considerably more indirect: The laws of physics provide a framework for processing information about nature. From this perspective physical models are mere tools that are partly discovered and partly designed with our own very human purposes in mind. This approach is decidedly pragmatic: when tools happen to be successful we do not say that they are *true*; we say that they are *useful*.

Is there any evidence in support of such an unorthodox view? The answer is yes. Indeed, if physics is an exercise in inference then we should expect it to include both ontic and epistemic concepts. The ontic concepts are meant to represent those entities in nature that are the subject of our interest.

They include the quantities, such as particle positions and field strengths, that we want to predict, to explain, and to control. The epistemic concepts, on the other hand, are the tools—the probabilities, the entropies, and the (information) geometries—that are used to carry out our inferences. The prominence of these epistemic elements strongly suggests that physics is not a mere mirror of nature; instead physics is an inference framework designed by humans for the purpose of facilitating their interactions with nature. The founders of quantum theory—Bohr, Heisenberg, Born, *etc.*—were quite aware of the epistemological and pragmatic elements in quantum mechanics (see e.g., [1]) but they wrote at a time when the tools of quantitative epistemology—the Bayesian and entropic methods of inference—had not yet been sufficiently developed.

Entropic Dynamics (ED) provides a framework for deriving dynamical laws as an application of entropic methods. (The principle of maximum entropy as a method for inference can be traced to E. T. Jaynes. For a pedagogical overview of Bayesian and entropic inference and further references see [2].) In ED, the dynamics is driven by entropy subject to the constraints appropriate to the problem at hand. It is through these constraints that the "physics" is introduced. Such a framework is extremely restrictive. For example, in order to adopt an epistemic view of the quantum state ψ it is not sufficient to merely assert that the probability $|\psi|^2$ represents a state of knowledge; this is a good start but it is not nearly enough. It is also necessary that the changes or updates of the epistemic ψ—which include *both* the unitary time evolution described the Schrödinger equation and the collapse of the wave function during measurement—be *derived* according to the established rules of inference. Therefore, in a truly entropic dynamics we are not allowed to postulate action principles that operate at some deeper sub-quantum level. Instead, the goal is to derive such action principles from entropy principles with suitably chosen constraints.

In this paper we collect and streamline results that have appeared in several publications (see [3–6] and references therein) to provide a self-contained overview of three types of ED: (1) standard diffusion; (2) Hamiltonian dynamics; and (3) quantum mechanics. First we tackle the case of a diffusion process which serves to address the central concern with the nature of time. In ED "entropic" time is a relational concept introduced as a book-keeping device designed to keep track of the accumulation of change. Time is constructed by (a) introducing the notion of "instants"; (b) showing that these instants are "ordered"; and (c) defining a measure of the interval that separates successive instants. The welcome new feature is that an arrow of time is generated automatically; entropic time is intrinsically directional.

The early formulation of ED [3] involved assumptions about auxiliary variables, the metric of configuration space, and the form of the quantum potential. All these assumptions have been subsequently removed. In [6] it was shown how the constraint that the dynamics be non-dissipative leads to a generic form of Hamiltonian dynamics, with its corresponding symplectic structure and action principle. Thus, in the context of ED action principles are *derived*; they are useful tools but they are not fundamental.

Different Hamiltonians lead to different dynamical laws. We show how considerations of information geometry provide the natural path to Hamiltonians that include the correct form of "quantum potential" and lead to the Schrödinger equation, and we also identify the constraints that describe motion in an external electromagnetic field.

Here, we focus on the derivation of the Schrödinger equation but the ED approach has been applied to several other topics in quantum mechanics that will not be reviewed here. These include the quantum measurement problem [5,7,8]; momentum, angular momentum, their uncertainty relations, and spin [9,10]; relativistic scalar fields [11]; the Bohmian limit [12]; and the extension to curved spaces [13].

There is vast literature on the attempts to reconstruct quantum mechanics and it is inevitable that the ED approach might resemble them in one aspect or another — after all, in order to claim success all these approaches must sooner or later converge to the same Schrödinger equation. However, there are important differences. For example, the central concern with the notion of time makes ED significantly different from other approaches that are also based on information theory (see e.g., [14–24]). ED also differs from those approaches that attempt to explain the emergence of quantum behavior as the effective statistical mechanics of some underlying sub-quantum dynamics which might possibly include some additional stochastic element (see e.g., [25–34]). Indeed, ED makes no reference to any sub-quantum dynamics whether classical, deterministic, or stochastic.

2. Entropic Dynamics

As with other applications of entropic methods, to derive dynamical laws we must first specify the microstates that are the subject of our inference — the subject matter — and then we must specify the prior probabilities and the constraints that represent the information that is relevant to our problem. (See e.g., [2].) We consider N particles living in a flat Euclidean space $\mathbf{X}$ with metric δ_{ab}. We assume that the particles have definite positions x_n^a and it is their unknown values that we wish to infer. (The index $n = 1 \ldots N$ denotes the particle and $a = 1, 2, 3$ the spatial coordinates.) For N particles the configuration space is $\mathbf{X}_N = \mathbf{X} \times \ldots \times \mathbf{X}$.

In this work ED is developed as a model for the quantum mechanics of particles. The same framework can be deployed to construct models for the quantum mechanics of fields, in which case it is the fields that are objectively "real" and have well-defined, albeit unknown, values [11].

The assumption that the particles have definite positions is in flat contradiction with the standard Copenhagen notion that quantum particles acquire definite positions only as a result of a measurement. For example, in the ED description of the double slit experiment we do not know which slit the quantum particle goes through but it most definitely goes through either one or the other.

We do not explain why motion happens but, given the information that it does, our task is to produce an estimate of what short steps we can reasonably expect. The next assumption is dynamical: we assume that the particles follow trajectories that are continuous. This means that displacements over finite distances can be analyzed as the accumulation of many infinitesimally short steps, and our first task is to find the transition probability density $P(x'|x)$ for a single short step from a given initial $x \in \mathbf{X}_N$ to an unknown neighboring $x' \in \mathbf{X}_N$. Later we will determine how such short steps accumulate to yield a finite displacement.

To find $P(x'|x)$ we maximize the (relative) entropy,

$$\mathcal{S}[P, Q] = -\int dx' \, P(x'|x) \log \frac{P(x'|x)}{Q(x'|x)} \tag{1}$$

To simplify the notation in all configuration space integrals we write $d^{3N}x = dx$, $d^{3N}x' = dx'$, and so on. $Q(x'|x)$ is the prior probability. It expresses our beliefs—or more precisely, the beliefs of an ideally rational agent—*before* any information about the motion is taken into account. The physically relevant information about the step is expressed in the form of constraints imposed on $P(x'|x)$—this is the stage at which the physics is introduced.

The prior—We adopt a prior $Q(x'|x)$ that represents a state of extreme ignorance: knowledge of the initial position x tells us nothing about x'. Such ignorance is expressed by assuming that $Q(x'|x)dx'$ is proportional to the volume element in $\mathbf{X}_N$. Since the space $\mathbf{X}_N$ is flat and a mere proportionality constant has no effect on the entropy maximization we can set $Q(x'|x) = 1$. The generalization to curved spaces is straightforward [13]. Another possible matter of concern is that the uniform prior is not normalizable and it is known that improper priors can sometimes be mathematically problematic. Fortunately, in our case no such difficulties arise. For microscopic particles any prior that is sufficiently flat over macroscopic scales turns out to lead to exactly the same physical predictions. We can, for example, use a Gaussian centered at x with a macroscopically large standard deviation and this leads to exactly the same transition probability.

The constraints—The first piece of information is that motion is continuous—motion consists of a succession of infinitesimally short steps. Each individual particle n will take a short step from x_n^a to $x_n'^a = x_n^a + \Delta x_n^a$ and we require that for each particle the expected squared displacement,

$$\langle \Delta x_n^a \Delta x_n^b \rangle \delta_{ab} = \kappa_n \qquad (n = 1 \ldots N) \tag{2}$$

takes some small value κ_n. Infinitesimally short steps are obtained by taking the limit $\kappa_n \to 0$. We will assume each κ_n to be independent of x to reflect the translational symmetry of $\mathbf{X}$. In order to describe non-identical particles, we assume that the value of κ_n depends on the particle index n.

The N constraints Equation (2) treat the particles as statistically independent and their accumulation eventually leads to a completely isotropic diffusion. But we know that particles can become correlated and even become entangled. We also know that motion is not normally isotropic; once particles are set in motion they tend to persist in it. This information is introduced through one additional constraint involving a "drift" potential $\phi(x)$ that is a function in configuration space, $x \in \mathbf{X}_N$. We impose that the expected displacements $\langle \Delta x_n^a \rangle$ along the direction of the gradient of ϕ satisfy

$$\langle \Delta x^A \rangle \partial_A \phi = \sum_{n=1}^{N} \langle \Delta x_n^a \rangle \frac{\partial \phi}{\partial x_n^a} = \kappa' \tag{3}$$

where capitalized indices such as $A = (n, a)$ include both the particle index and its spatial coordinate; $\partial_A = \partial/\partial x^A = \partial/\partial x_n^a$; and κ' is another small, but for now unspecified, position-independent constant.

The introduction of the drift potential ϕ will not be justified at this point. The idea is that we can make progress by identifying the constraints even when their physical origin remains unexplained. This situation is not unlike classical mechanics, where identifying the forces is useful even in situations where their microscopic origin is not understood. We do however make two brief comments. First, in section 9 we shall see that ED in an external electromagnetic field is described by constraints that are formally similar to Equation (3). There we shall show that the effects of the drift potential ϕ and the electromagnetic vector potential A_a are intimately related—a manifestation of gauge

symmetry—suggesting that whatever ϕ might be, it is as "real" as A_a. The second comment is that elsewhere, in the context of a particle with spin, we will see that the drift potential can be given a natural geometric interpretation as an angular variable. This imposes the additional condition that the integral of ϕ over any closed loop is quantized, $\oint d\phi = 2\pi\nu$ where ν is an integer.

Maximizing $\mathcal{S}[P,Q]$ in Equation (1) subject to the $N+2$ constraints Equations (2) and (3) plus normalization yields a Gaussian distribution,

$$P(x'|x) = \frac{1}{\zeta} \exp \sum_n \left[-\frac{1}{2}\alpha_n \, \Delta x_n^a \Delta x_n^b \delta_{ab} + \alpha' \Delta x_n^a \frac{\partial \phi}{\partial x_n^a} \right] \tag{4}$$

where $\zeta = \zeta(x, \alpha_n, \alpha')$ is a normalization constant and the Lagrange multipliers α_n and α' are determined from

$$\partial \log \zeta / \partial \alpha_n = -\kappa_n/2 \quad \text{and} \quad \partial \log \zeta / \partial \alpha' = \kappa' \tag{5}$$

The distribution $P(x'|x)$ is conveniently rewritten as

$$P(x'|x) = \frac{1}{Z} \exp \left[-\frac{1}{2}\sum_n \alpha_n \, \delta_{ab}(\Delta x_n^a - \langle \Delta x_n^a \rangle)(\Delta x_n^b - \langle \Delta x_n^b \rangle) \right] \tag{6}$$

where Z is a new normalization constant. A generic displacement $\Delta x_n^a = x_n'^a - x_n^a$ can be expressed as an expected drift plus a fluctuation,

$$\Delta x_n^a = \langle \Delta x_n^a \rangle + \Delta w_n^a \, , \quad \text{where} \quad \langle \Delta x_n^a \rangle = \frac{\alpha'}{\alpha_n}\delta^{ab}\frac{\partial \phi}{\partial x_n^b} \tag{7}$$

$$\langle \Delta w_n^a \rangle = 0 \quad \text{and} \quad \langle \Delta w_n^a \Delta w_n^b \rangle = \frac{1}{\alpha_n}\delta^{ab} \tag{8}$$

From these equations we can get a first glimpse into the meaning of the multipliers α_n and α'. For very short steps, as $\alpha_n \to \infty$, the fluctuations become dominant: the drift is $\langle \Delta x_n^a \rangle \sim O(\alpha_n^{-1})$ while $\Delta w_n^a \sim O(\alpha_n^{-1/2})$. This implies that, as in Brownian motion, the trajectory is continuous but not differentiable. In the ED approach a particle has a definite position, but its velocity—the tangent to the trajectory—is completely undefined. We can also see that the effect of α' is to enhance or suppress the magnitude of the drift relative to the fluctuations — a subject that is discussed in detail in [12]. However, for our current purposes we can absorb α' into the so far unspecified drift potential, $\alpha'\phi \to \phi$, which amounts to setting $\alpha' = 1$.

3. Entropic Time

The goal is to derive dynamical laws as an application of inference methods, although the latter make no reference to time. Therefore, additional assumptions are needed to specify what constitutes "time" in our framework.

The foundation to any notion of time is dynamics. We must first identify a notion of "instant" that properly takes into account the inferential nature of entropic dynamics. Time is then is constructed as a device to keep track of change or, more explicitly, of the accumulation of repeated small changes.

3.1. Time as an Ordered Sequence of Instants

Entropic dynamics is generated by the short-step transition probability $P(x'|x)$, Equation (6). In the ith step, which takes the system from $x = x_{i-1}$ to $x' = x_i$, both x_{i-1} and x_i are unknown. Integrating the joint probability, $P(x_i, x_{i-1})$, over x_{i-1} gives

$$P(x_i) = \int dx_{i-1} P(x_i, x_{i-1}) = \int dx_{i-1} P(x_i|x_{i-1}) P(x_{i-1}) \tag{9}$$

These equations involve no assumptions—they are true by virtue of the laws of probability. However, if $P(x_{i-1})$ happens to be the probability of different values of x_{i-1} *at an "instant" labelled t*, then we can interpret $P(x_i)$ as the probability of values of x_i *at the next "instant"* which we will label t'. Accordingly, we write $P(x_{i-1}) = \rho(x, t)$ and $P(x_i) = \rho(x', t')$ so that

$$\rho(x', t') = \int dx \, P(x'|x) \rho(x, t) \tag{10}$$

Nothing in the laws of probability leading to Equation (9) forces the interpretation Equation (10) on us—this is an independent assumption about what constitutes time in our model. We use Equation (10) to define what we mean by an instant: *if the distribution $\rho(x, t)$ refers to one instant t, then the distribution $\rho(x', t')$ generated by $P(x'|x)$ defines what we mean by the "next" instant t'*. The iteration of this process defines the dynamics: entropic time is constructed instant by instant: $\rho(t')$ is constructed from $\rho(t)$, $\rho(t'')$ is constructed from $\rho(t')$, and so on.

The inferential nature of the construction can be phrased more explicitly. Once we have decided on the relevant information necessary for predicting future behavior—the distributions $\rho(x, t)$ and $P(x'|x)$—we can say that this information defines what we mean by an "instant". Furthermore, Equation (10) shows that:

> *Time is* designed *in such a way that given the present the future is independent of the past.*

An equation such as Equation (10) is commonly employed to define Markovian behavior, in which case it is sometimes known as the Chapman–Kolmogorov equation. Markovian processes are such that once an external notion of time is given, defined perhaps by an external clock, the specification of the state of the system at time t is sufficient to fully determine its state after time t—no additional information about times prior to t is needed. It should be emphasized that we do not make a Markovian assumption. We are concerned with a different problem: we do not use Equation (10) to define a Markovian process *in an already existing background time*; we use it to construct time itself.

3.2. The Arrow of Entropic Time

The notion of time constructed according to Equation (10) incorporates an intrinsic directionality: there is an absolute sense in which $\rho(x, t)$ is prior and $\rho(x', t')$ is posterior. To construct the time-reversed evolution we just write

$$\rho(x, t) = \int dx' \, P(x|x') \rho(x', t') , \tag{11}$$

where according to the rules of probability theory $P(x|x')$ is related to $P(x'|x)$ in Equation (6) by Bayes' theorem,

$$P(x|x') = \frac{\rho(x, t)}{\rho(x', t')} P(x'|x) . \tag{12}$$

Note that this is not a mere exchange of primed and unprimed quantities: $P(x'|x)$ is the Gaussian distribution, Equation (6), obtained using the maximum entropy method; in contrast, the time-reversed $P(x|x')$ obtained using Bayes' theorem, Equation (12), will not in general be Gaussian. The asymmetry between the inferential past and the inferential future is traced to the asymmetry between priors and posteriors.

The subject of the arrow of time has a vast literature (see e.g., [35,36]). The puzzle has been how to explain the asymmetric arrow from underlying symmetric laws of nature. The solution offered by ED is that there are no symmetric underlying laws and the asymmetry is the inevitable consequence of entropic inference. From the point of view of ED, the challenge is not to explain the arrow of time, but rather the reverse, how to explain the emergence of symmetric laws within an intrinsically asymmetric entropic framework. As we shall see below the derived laws of physics—e.g., the Schrödinger equation—are time-reversible but *entropic time itself only flows forward.*

3.3. Duration: A Convenient Time Scale

To complete the model of entropic time we need to specify the interval Δt between successive instants. The basic criterion is that *duration is defined so that motion looks simple.* From Equations (7) and (8) for short steps (large α_n) the motion is dominated by the fluctuations Δw_n. Therefore specifying Δt amounts to specifying the multipliers α_n in terms of Δt.

To start, we appeal to symmetry: in order that the fluctuations $\langle \Delta w_n^a \Delta w_n^b \rangle$ reflect the symmetry of translations in space and time—a time that, just like Newton's, flows "equably everywhere and everywhen"—we choose α_n to be independent of x and t, and we choose Δt so that $\alpha_n \propto 1/\Delta t$. More explicitly, we write

$$\alpha_n = \frac{m_n}{\eta} \frac{1}{\Delta t} \tag{13}$$

where the proportionality constants have been written in terms of some particle-specific constants m_n, which will eventually be identified with the particle masses, and an overall constant η that fixes the units of the m_ns relative to the units of time and will later be regraduated into $\hbar$.

Before discussing the implications of the choice Equation (13) it is useful to consider the geometry of the N-particle configuration space, $\mathbf{X}_N$.

4. The Information Metric of Configuration Space

We have assumed that the geometry of the single particle spaces $\mathbf{X}$ is described by the Euclidean metric δ_{ab}. We can expect that the N-particle configuration space, $\mathbf{X}_N = \mathbf{X} \times \ldots \times \mathbf{X}$, will also have a flat geometry, but the relative contribution of different particles to the metric remains undetermined. Should very massive particles contribute the same as very light particles? The answer is provided by information geometry.

To each point $x \in \mathbf{X}_N$ there corresponds a probability distribution $P(x'|x)$. Therefore $\mathbf{X}_N$ is a statistical manifold and up to an arbitrary global scale factor its geometry is uniquely determined by the information metric,

$$\gamma_{AB} = C \int dx' \, P(x'|x) \frac{\partial \log P(x'|x)}{\partial x^A} \frac{\partial \log P(x'|x)}{\partial x^B} \tag{14}$$

where C is an arbitrary positive constant (see e.g., [2]). A straightforward substitution of Equations (6) and (13) into Equation (14) in the limit of short steps ($\alpha_n \to \infty$) yields

$$\gamma_{AB} = \frac{Cm_n}{\eta\Delta t}\delta_{nn'}\,\delta_{ab} = \frac{Cm_n}{\eta\Delta t}\delta_{AB} \tag{15}$$

We see that γ_{AB} diverges as $\Delta t \to 0$. The reason for this is not hard to find. As the Gaussian distributions $P(x'|x)$ and $P(x'|x + \Delta x)$ become more sharply peaked and it is easier to distinguish one from the other which translates into a greater information distance, $\gamma_{AB} \to \infty$. In order to define a distance that remains meaningful for arbitrarily small Δt it is convenient to choose $C \propto \Delta t$. In what follows the metric tensor will always appear in combinations such as $\gamma_{AB}\Delta t/C$. It is therefore convenient to define the "mass" tensor,

$$m_{AB} = \frac{\eta\Delta t}{C}\gamma_{AB} = m_n\delta_{AB} \tag{16}$$

Its inverse,

$$m^{AB} = \frac{C}{\eta\Delta t}\gamma^{AB} = \frac{1}{m_n}\delta^{AB} \tag{17}$$

is called the "diffusion" tensor.

We can now summarize our results so far. The choice Equation (13) of the multipliers α_n simplifies the dynamics: $P(x'|x)$ in Equation (6) is a standard Wiener process. A generic displacement, Equation (7), is

$$\Delta x^A = b^A\Delta t + \Delta w^A \tag{18}$$

where $b^A(x)$ is the drift velocity,

$$\langle\Delta x^A\rangle = b^A\Delta t \quad \text{with} \quad b^A = \frac{\eta}{m_n}\delta^{AB}\partial_B\phi = \eta m^{AB}\partial_B\phi \tag{19}$$

and the fluctuations Δw^A are such that

$$\langle\Delta w^A\rangle = 0 \quad \text{and} \quad \langle\Delta w^A\Delta w^B\rangle = \frac{\eta}{m_n}\delta^{AB}\Delta t = \eta m^{AB}\Delta t \tag{20}$$

We are now ready to comment on the implications of the choice of time scale Δt and of multipliers α_n, Equation (13). The first remark is on the nature of clocks: In Newtonian mechanics the prototype of a clock is the free particle, and time is defined so as to simplify the motion of free particles—they move equal distances in equal times. In ED the prototype of a clock is a free particle too—for sufficiently short times all particles are free—and time here is also defined to simplify the description of their motion: the particle undergoes equal fluctuations in equal times.

The second remark is on the nature of mass. The particle-specific constants m_n will, in due course, be called "mass" and Equation (20) provides their interpretation: mass is an inverse measure of fluctuations. Thus, up to overall constants the metric of configuration space is the mass tensor and its inverse is the diffusion tensor. In standard QM there are two mysteries: "Why quantum fluctuations?" and "What is mass?" ED offers some progress in that instead of two mysteries there is only one. Fluctuations and mass are two sides of the same coin.

Finally we note the formal similarity to Nelson's stochastic mechanics [25]. The similarity is to be expected—all theories that converge on the Schrödinger equation must at some point become formally

similar—but our epistemic interpretation differs radically from Nelson's ontic interpretation and avoids the difficulties discussed in [37].

5. Diffusive Dynamics

Equation (10) is the dynamical equation for the evolution of $\rho(x, t)$. It is written in integral form but it can be written in differential form as a Fokker–Planck (FP) equation (see e.g., [2])

$$\partial_t \rho = -\partial_A \left(b^A \rho \right) + \frac{1}{2} \eta m^{AB} \partial_A \partial_B \rho \tag{21}$$

or equivalently as a continuity equation,

$$\partial_t \rho = -\partial_A \left(\rho v^A \right) \tag{22}$$

where

$$v^A = b^A + u^A \tag{23}$$

is the velocity of the probability flow or *current velocity* and

$$u^A = -\eta m^{AB} \partial_B \log \rho^{1/2} \tag{24}$$

is called the *osmotic velocity*—it represents the tendency for probability to flow down the density gradient. Since both b^A and u^A are gradients, it follows that the current velocity is a gradient too,

$$v^A = m^{AB} \partial_B \Phi \quad \text{where} \quad \Phi = \eta(\phi - \log \rho^{1/2}) . \tag{25}$$

The FP equation,

$$\partial_t \rho = -\partial_A \left(\rho m^{AB} \partial_B \Phi \right) , \tag{26}$$

can be conveniently rewritten in the alternative form

$$\partial_t \rho = \frac{\delta \tilde{H}}{\delta \Phi} \tag{27}$$

for some suitably chosen functional $\tilde{H}[\rho, \Phi]$. It is easy to check that the appropriate functional $\tilde{H}$ is

$$\tilde{H}[\rho, \Phi] = \int dx \, \frac{1}{2} \rho m^{AB} \partial_A \Phi \partial_B \Phi + F[\rho] \tag{28}$$

where the integration constant $F[\rho]$ is some unspecified functional of ρ.

With these results we have demonstrated that a specific form of dynamics—a standard diffusion process—can be derived from principles of entropic inference. This diffusive dynamics can be written in different but equivalent ways—Equations (21), (22), (26) and (27) are all equivalent. Next we turn our attention to other forms of dynamics, such as quantum or classical mechanics which require a somewhat different choice of constraints.

6. Hamiltonian Dynamics

The previous discussion has led us to a *standard* diffusion, in which the density ρ evolves under the influence of some externally fixed drift potential ϕ. However, in quantum dynamics we require a second degree of freedom, the phase of the wave function. The extra degree of freedom is introduced into ED by replacing the constraint of a fixed drift potential ϕ by an evolving constraint in which at each time step the potential ϕ is readjusted in response to the evolving ρ.

To find the appropriate readjustment of ϕ we borrow an idea of Nelson's [38] and impose that the potential ϕ be updated in such a way that a certain functional, later called "energy", remains constant. The next challenge is to identify the appropriate functional form of this energy, but before this we make two remarks.

The standard procedure in mechanics is to derive the conservation of energy from the invariance of the action under time translations, but here we do not have an action yet. The logic of our derivation runs in the opposite direction: we first identify the conservation of an energy as the piece of information that is relevant to our inferences, and from it we derive Hamilton's equations and their associated action principle.

Imposing energy conservation appears to be natural because it agrees with our classical preconceptions of what mechanics is like. But ED is not at all like classical mechanics. Indeed, Equation (18) is the kind of equation (a Langevin equation) that characterizes a Brownian motion in the limit of *infinite* friction. Therefore in the ED approach to quantum theory particles seem to be subject to infinite friction while suffering zero dissipation. Such a strange dynamics can hardly be called "mechanics", much less "classical".

The Ensemble Hamiltonian—The energy functional that codifies the correct constraint is of the form Equation (28). We therefore impose that, irrespective of the initial conditions, the potential ϕ will be updated in such a way that the functional $\tilde{H}[\rho, \Phi]$ in Equation (28) is always conserved,

$$\frac{d\tilde{H}}{dt} = \int dx \left[\frac{\delta \tilde{H}}{\delta \Phi} \partial_t \Phi + \frac{\delta \tilde{H}}{\delta \rho} \partial_t \rho \right] = 0 \tag{29}$$

Using Equation (27) we get

$$\frac{d\tilde{H}}{dt} = \int dx \left[\partial_t \Phi + \frac{\delta \tilde{H}}{\delta \rho} \right] \partial_t \rho = 0 \tag{30}$$

We require that $\tilde{H} = $ const. for arbitrary choices of the initial values of ρ and Φ. From Equation (26) we see that this amounts to imposing $d\tilde{H}/dt = 0$ for arbitrary choices of $\partial_t \rho$. Therefore the requirement that $\tilde{H}$ be conserved for arbitrary initial conditions amounts to imposing that

$$\partial_t \Phi = -\frac{\delta \tilde{H}}{\delta \rho} \tag{31}$$

Equations (27) and (31) have the form of a canonically conjugate pair of Hamilton's equations. The field ρ is a generalized coordinate and Φ is its canonical momentum. The conserved functional $\tilde{H}[\rho, \Phi]$ in Equation (28) will be called the ensemble Hamiltonian. We conclude that non-dissipative ED leads to Hamiltonian dynamics.

Equation (31) leads to a generalized Hamilton–Jacobi equation,

$$\partial_t \Phi = -\frac{1}{2} m^{AB} \partial_A \Phi \partial_B \Phi - \frac{\delta F}{\delta \rho} \tag{32}$$

The Action, Poisson Brackets, *etc.*—Now that we have Hamilton's equations, (27) and (31), we can invert the usual procedure and *construct* an action principle from which they can be derived. Define the differential

$$\delta A = \int dt \int dx \left[\left(\partial_t \rho - \frac{\delta \tilde{H}}{\delta \Phi} \right) \delta \Phi - \left(\partial_t \Phi + \frac{\delta \tilde{H}}{\delta \rho} \right) \delta \rho \right] \tag{33}$$

and then integrate to get the action

$$A[\rho, \Phi] = \int dt \left(\int dx \, \Phi \dot{\rho} - \tilde{H}[\rho, \Phi] \right) \tag{34}$$

By construction, imposing $\delta A = 0$ leads to Equations (27) and (31).

The time evolution of any arbitrary functional $f[\rho, \Phi]$ is given by a Poisson bracket,

$$\frac{d}{dt} f[\rho, \Phi] = \int dx \left[\frac{\delta f}{\delta \rho} \frac{\delta \tilde{H}}{\delta \Phi} - \frac{\delta f}{\delta \Phi} \frac{\delta \tilde{H}}{\delta \rho} \right] = \{f, \tilde{H}\} \tag{35}$$

which shows that the ensemble Hamiltonian $\tilde{H}$ is the generator of time evolution. Similarly, under a spatial displacement ε^a the change in $f[\rho, \Phi]$ is

$$\delta_\varepsilon f[\rho, \Phi] = \left\{ f, \tilde{P}_a \varepsilon^a \right\} \tag{36}$$

where

$$\tilde{P}_a = \int d^{3N}x \, \rho \sum_n \frac{\partial \Phi}{\partial x_n^a} = \int d^{3N}x \, \rho \frac{\partial \Phi}{\partial X^a} \tag{37}$$

is interpreted as the expectation of the total momentum, and X^a are the coordinates of the center of mass,

$$X^a = \frac{1}{M} \sum_n m_n x_n^a \tag{38}$$

A Schrödinger-Like Equation—We can always combine ρ and Φ to define the family of complex functions,

$$\Psi_k = \rho^{1/2} \exp(ik\Phi/\eta) \tag{39}$$

where k is some arbitrary positive constant. Then the two coupled Equations (27) and (31) can be written as a single complex Schrödinger-like equation,

$$i \frac{\eta}{k} \partial_t \Psi_k = -\frac{1}{2} \frac{\eta^2}{k^2} m^{AB} \partial_A \partial_B \Psi_k + \frac{1}{2} \frac{\eta^2}{k^2} m^{AB} \frac{\partial_A \partial_B |\Psi_k|}{|\Psi_k|} \Psi_k + \frac{\delta F}{\delta \rho} \Psi_k \tag{40}$$

The reason for the parameter k will become clear shortly, but even at this stage we can already anticipate that η/k will play the role of $\hbar$.

7. Information Geometry and the Quantum Potential

Different choices of the functional $F[\rho]$ in Equation (28) lead to different dynamics. Earlier we invoked information geometry, Equation (14), to define the metric m_{AB} induced in configuration space by the transition probabilities $P(x'|x)$. To motivate the particular choice of the functional $F[\rho]$ that leads to quantum theory we appeal to information geometry once again.

Consider the family of distributions $\rho(x|\theta)$ that are generated from a distribution $\rho(x)$ by pure translations by a vector θ^A, $\rho(x|\theta) = \rho(x - \theta)$. The extent to which $\rho(x|\theta)$ can be distinguished from the slightly displaced $\rho(x|\theta + d\theta)$ or, equivalently, the information distance between θ^A and $\theta^A + d\theta^A$, is given by

$$d\ell^2 = g_{AB}d\theta^A d\theta^B \tag{41}$$

where

$$g_{AB}(\theta) = \int dx \frac{1}{\rho(x - \theta)} \frac{\partial \rho(x - \theta)}{\partial \theta^A} \frac{\partial \rho(x - \theta)}{\partial \theta^B} \tag{42}$$

Changing variables $x - \theta \to x$ yields

$$g_{AB}(\theta) = \int dx \frac{1}{\rho(x)} \frac{\partial \rho(x)}{\partial x^A} \frac{\partial \rho(x)}{\partial x^B} = I_{AB}[\rho] \tag{43}$$

The Functional $F[\rho]$—The simplest choice of functional $F[\rho]$ is linear in ρ, $F[\rho] = \int dx\, \rho V$, where $V(x)$ is some function that will be recognized as the familiar scalar potential. Since ED aims to derive the laws of physics from a framework for inference, it is natural to expect that the Hamiltonian might also contain terms that are of a purely informational nature. We have identified two such tensors: one is the information metric of configuration space $\gamma_{AB} \propto m_{AB}$, another is $I_{AB}[\rho]$. The simplest nontrivial scalar that can be constructed from them is the trace $m^{AB}I_{AB}$. This suggests

$$F[\rho] = \xi m^{AB} I_{AB}[\rho] + \int dx\, \rho V \tag{44}$$

where $\xi > 0$ is a constant that controls the relative strength of the two contributions. The term $m^{AB}I_{AB}$ is sometimes called the "quantum" or the "osmotic" potential. This relation between the quantum potential and the Fisher information was pointed out in [39]. From Equation (43) we see that $m^{AB}I_{AB}$ is a contribution to the energy such that those states that are more smoothly spread out tend to have lower energy. The case $\xi < 0$ leads to instabilities and is therefore excluded; the case $\xi = 0$ leads to a qualitatively different theory and will be discussed elsewhere [12].

With this choice of $F[\rho]$ the generalized Hamilton-Jacobi Equation (32) becomes

$$-\partial_t \Phi = \frac{1}{2} m^{AB} \partial_A \Phi \partial_B \Phi + V - 4\xi m^{AB} \frac{\partial_A \partial_B \rho^{1/2}}{\rho^{1/2}} \tag{45}$$

8. The Schrödinger Equation

Substituting Equation (44) into Equation (40) gives a Schrödinger-like equation,

$$i\frac{\eta}{k} \partial_t \Psi_k = -\frac{\eta^2}{2k^2} m^{AB} \partial_A \partial_B \Psi_k + V\Psi_k + \left(\frac{\eta^2}{2k^2} - 4\xi\right) m^{AB} \frac{\partial_A \partial_B |\Psi_k|}{|\Psi_k|} \Psi_k \tag{46}$$

the beauty of which is severely marred by the non-linear last term.

Regraduation—We can now make good use of the freedom afforded by the arbitrary constant k. Since the physics is fully described by ρ and Φ, the different choices of k in Ψ_k all describe the same theory. Among all these equivalent descriptions, it is clearly to our advantage to pick the k that is most *convenient*—a process usually known as "regraduation". Other notable examples of regraduation include the Kelvin choice of absolute temperature, the Cox derivation of the sum and product rule for probabilities, and the derivation of the sum and product rules for quantum amplitudes [2,15].

A quick examination of Equation (46) shows that the optimal k is such that the non-linear term drops out. The optimal choice, which we denote $\hat{k}$, is

$$\hat{k} = (\frac{\eta^2}{8\xi})^{1/2} \tag{47}$$

We can now identify the optimal regraduated $\eta/\hat{k}$ with Planck's constant $\hbar$,

$$\frac{\eta}{\hat{k}} = (8\xi)^{1/2} = \hbar \tag{48}$$

and Equation (46) becomes the linear Schrödinger equation,

$$i\hbar\partial_t\Psi = -\frac{\hbar^2}{2}m^{AB}\partial_A\partial_B\Psi + V\Psi = \sum_n \frac{-\hbar^2}{2m_n}\nabla_n^2\Psi + V\Psi \tag{49}$$

where the wave function is $\Psi = \rho e^{i\Phi/\hbar}$. The constant $\xi = \hbar^2/8$ in Equation (44) turns out to be crucial: it defines the numerical value of what we call Planck's constant and sets the scale that separates quantum from classical regimes.

The conclusion is that for any positive value of the constant ξ it is always possible to regraduate Ψ_k to a physically equivalent but more convenient description where the Schrödinger equation is linear. From the ED perspective the linear superposition principle and the complex Hilbert spaces are important because they are convenient, but not because they are fundamental.

9. ED in an External Electromagnetic Field

In ED the information that is physically relevant for prediction is codified into constraints that reflect that motion is (a) continuous, (b) correlated and directional, and (c) non-dissipative. These constraints are expressed by Equations (2), (3), and (31) respectively. In this section we show that interactions can be introduced by imposing additional constraints.

As an explicit illustration, we show that the effect of an external electromagnetic field is modelled by a constraint on the component of displacements along a certain direction represented by the vector potential $A_a(x)$. For each particle n we impose the constraint

$$\langle \Delta x^a \rangle A_a(x_n) = \kappa_n'', \quad (n = 1 \ldots N) \tag{50}$$

where κ_n'' is a particle-dependent constant that reflects the strength of the coupling to A_a.

The resemblance between Equation (50) and the drift potential constraint, Equation (3), is very significant—as we shall see shortly it leads to gauge symmetries—but there also are significant

differences. Note that Equation (3) is a single constraint acting in the N-particle configuration space—it involves the drift potential $\phi(x)$ with $x \in \mathbf{X}_N$. In contrast, Equation (50) is N constraints acting in the 1-particle space—the vector potential $A_a(x_n)$ is a function in 3D space, $x_n \in \mathbf{X}$.

Except for minor changes the development of ED proceeds as before. The transition probability $P(x'|x)$ that maximizes the entropy $\mathcal{S}[P,Q]$, Equation (1), subject to Equations (2), (3) and (50) and normalization, is

$$P(x'|x) = \frac{1}{\zeta} \exp - \sum_n \left[\frac{1}{2} \alpha_n \delta_{ab} \Delta x_n^a \Delta x_n^b - \left(\alpha' \frac{\partial \phi}{\partial x_n^a} - \alpha_n'' A_a(x_n) \right) \Delta x_n^a \right] \tag{51}$$

which includes an additional set of Lagrange multipliers α_n''. Next use Equations (13) and (16) to get

$$P(x'|x) = \frac{1}{\zeta} \exp - \left[\frac{1}{2\eta \Delta t} m_{AB} \Delta x^A \Delta x^B - \left(\partial_A \phi - \frac{1}{\eta} A_A \right) \Delta x^A \right] \tag{52}$$

where we have absorbed α' into ϕ, $\alpha' \phi \to \phi$, and written

$$A_A(x) = \eta \alpha_n'' A_a(x_n) \tag{53}$$

as a vector in configuration space. As in Equation (18), a generic displacement Δx^A can be expressed in terms of a expected drift plus a fluctuation, $\Delta x^A = b^A \Delta t + \Delta w^A$, but the drift velocity now includes a new term,

$$b^A = m^{AB} \left(\eta \partial_A \phi - A_A \right) \tag{54}$$

The fluctuations Δw^A, Equation (20), remain unchanged.

A very significant feature of the transition probability $P(x'|x)$ is its invariance under gauge transformations,

$$\frac{\partial \phi}{\partial x_n^a} \to \frac{\partial \tilde{\phi}}{\partial x_n^a} = \frac{\partial \phi}{\partial x_n^a} + \frac{1}{\alpha'} \frac{\partial \chi(x_n)}{\partial x_n^a} \tag{55}$$

$$A_a(x_n) \to \tilde{A}_a(x_n) = A_a(x_n) + \frac{1}{\alpha''} \frac{\partial \chi(x_n)}{\partial x_n^a} \tag{56}$$

Note that these transformations are local in space. (The vector potential $A_a(x_n)$ and the gauge function $\chi(x_n)$ are functions in space.) They can be written in the N-particle configuration space,

$$\phi \to \tilde{\phi} = \phi + \frac{1}{\eta} \bar{\chi} \tag{57}$$

$$A_A \to \tilde{A}_A = A_A + \partial_A \bar{\chi} \tag{58}$$

where

$$\bar{\chi}(x) = \sum_n \chi(x_n) \tag{59}$$

The accumulation of many small steps is described by a Fokker–Planck equation which can be written either as a continuity equation, Equation (22), or in its Hamiltonian form, Equation (27). As might be expected, the current velocity v^A, Equation (25), and the ensemble Hamiltonian $\tilde{H}$, Equation (28), must be suitably modified,

$$v^A = m^{AB} \left(\partial_B \Phi - A_B \right) \quad \text{with} \quad \Phi = \eta(\phi - \log \rho^{1/2}) \tag{60}$$

and

$$\tilde{H}[\rho, \Phi] = \int dx \left[\frac{1}{2} \rho m^{AB} \left(\partial_A \Phi - A_A \right) \left(\partial_B \Phi - A_B \right) + \rho V + \frac{\hbar^2}{8\rho} m^{AB} \partial_A \rho \partial_B \rho \right] \tag{61}$$

As a shortcut here, we have adopted the same functional $F[\rho]$ motivated by information-geometry, Equation (44), and set $\xi = \hbar^2/8$. The new FP equation now reads,

$$\partial_t \rho = -\partial_A \left[\rho m^{AB} \left(\partial_B \Phi - A_B \right) \right] \tag{62}$$

The requirement that $\tilde{H}$ be conserved for arbitrary initial conditions amounts to imposing the second Hamilton equation, Equation (31), which leads to the Hamilton–Jacobi equation,

$$\partial_t \Phi = -\frac{1}{2} m^{AB} \left(\partial_A \Phi - A_A \right) \left(\partial_B \Phi - A_B \right) + V - \frac{\hbar^2}{2} m^{AB} \frac{\partial_A \partial_B \rho^{1/2}}{\rho^{1/2}} \tag{63}$$

Finally, we combine ρ and Φ into a single wave function, $\Psi = \rho e^{i\Phi/\hbar}$, to obtain the Schrödinger equation,

$$i\hbar \partial_t \Psi_k = -\frac{\hbar^2}{2} m^{AB} (\partial_A - \frac{i}{\hbar} A_A)(\partial_B - \frac{i}{\hbar} A_B) \Psi_k + V \Psi_k \tag{64}$$

We conclude with two comments. First, the covariant derivative

$$D_A = \partial_A - \frac{i}{\hbar} A_A = \frac{\partial}{\partial x_n^a} - \frac{i}{\hbar} \eta \alpha_n'' A_a(x_n) \tag{65}$$

in Equation (64) can be written in the standard notation,

$$D_A = \frac{\partial}{\partial x_n^a} - \frac{ie_n}{\hbar c} A_a(x_n) \tag{66}$$

where e_n is the electric charge of particle n in units where c is the speed of light. Comparing Equation (65) with Equation (66) allows us to interpret the Lagrange multipliers

$$\alpha_n'' = \frac{e_n}{\eta c} \tag{67}$$

in terms of the electric charges e_n and the speed of light c. Thus, in ED electric charge e_n is essentially a Lagrange multiplier α_n'' that regulates the response to the external electromagnetic potential A_a.

The second comment is that the derivation above is limited to static external potentials, $\partial_t V = 0$ and $\partial_t A_a = 0$, so that energy is conserved. This limitation is easily lifted. For time-dependent potentials the relevant energy condition must take into account the work done by external sources: we require that the energy increase at the rate

$$\frac{d\tilde{H}}{dt} = \frac{\partial \tilde{H}}{\partial t} \tag{68}$$

The net result is that Equations (62)–(64) remain valid for time-dependent external potentials.

10. Some Remarks and Conclusions

Are there any new predictions?—Our goal has been to derive dynamical laws, and in particular quantum theory, as an example of entropic inference. This means that, to the extent that we succeed and derive quantum mechanics and not some other theory, we should not expect predictions that deviate from those of the standard quantum theory — at least not in the nonrelativistic regime discussed here. However, the motivation behind the ED program lies in the conviction that it will eventually allow us to extend it to other realms, such as gravity or cosmology, where the status of quantum theory is more questionable.

The Wallstrom objection—An important remaining question is whether the Fokker–Planck and the generalized Hamilton–Jacobi equations, Equations (27) and (31), are fully equivalent to the Schrödinger equation. This point was first raised by Wallstrom [40,41] in the context of Nelson's stochastic mechanics [25] and concerns the single- or multi-valuedness of phases and wave functions. Briefly, the objection is that stochastic mechanics leads to phases Φ and wave functions Ψ that are either both multi-valued or both single-valued. Both alternatives are unsatisfactory: quantum mechanics forbids multi-valued wave functions, while single-valued phases can exclude physically relevant states (e.g., states with non-zero angular momentum). Here we do not discuss this issue in any detail except to note that the objection does not apply once particle spin is incorporated into ED. As shown by Takabayasi [42], a similar result holds for the hydrodynamical formalism. The basic idea is that, as mentioned earlier, the drift potential ϕ should be interpreted as an angle. Then the integral of the phase $d\Phi$ over a closed path gives precisely the quantization condition that guarantees that wave functions remain single-valued even for multi-valued phases.

Epistemology *vs*. ontology—Dynamical laws have been derived as an example of entropic dynamics. In this model "reality" is reflected in the positions of the particles and our "limited information about reality" is represented in the probabilities as they are updated to reflect the physically relevant constraints.

Quantum non-locality— ED may appear classical because no quantum probabilities were introduced. But this is not so. Probabilities, in this approach, are neither classical nor quantum; they are tools for inference. Phenomena that would normally be considered non-classical, such as non-local correlations, are the natural result of including the quantum potential term in the ensemble Hamiltonian.

On dynamical laws—Action principles are not fundamental; they are convenient ways to summarize the dynamical laws derived from the deeper principles of entropic inference. The requirement that an energy be conserved is an important piece of information (*i.e.*, a constraint) which will probably receive its full justification once a completely relativistic extension of entropic dynamics to gravity is developed.

On entropic *vs*. physical time—The derivation of laws of physics as examples of inference led us to introduce the informationally motivated notion of entropic time, which includes assumptions about the concepts of instant, simultaneity, ordering, and duration. It is clear that entropic time is useful, but is this the actual, real, "physical" time? The answer is yes. By deriving the Schrödinger equation (from which we can obtain the classical limit) we have shown that the t that appears in the laws of physics is entropic time. Since these are the equations that we routinely use to design and calibrate our clocks, we conclude that *what clocks measure is entropic time*. No notion of time that is in any way deeper or more "physical" is needed. Most interestingly, the entropic model automatically includes an arrow of time.

Acknowledgments

Conflicts of Interest

The author declares no conflict of interest.

References

1. Stapp, H.P. The Copenhagen Interpretation. *Am. J. Phys.* **1972**, *40*, 1098–1116.
2. Caticha, A. *Entropic Inference and the Foundations of Physics*; USP Press: São Paulo, Brazil, 2012; Available online: http://www.albany.edu/physics/ACaticha-EIFP-book.pdf (accessed on 26 August 2015).
3. Caticha, A. Entropic dynamics, time, and quantum theory. *J. Phys. A Math. Theor.* **2011**, *44*, 225303.
4. Caticha, A. Entropic time. In Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Chamonix, France, 4–9 July 2010; Mohammad-Djafari, A., Ed.; AIP Publishing: Melville, NY, USA, 2010; Volume 1305, pp. 200–212.
5. Caticha, A. Entropic dynamics: An inference approach to quantum theory, time and measurement. *J. Phys. Conf. Ser.* **2014**, *504*, 012009.
6. Caticha, A.; Bartolomeo, D.; Reginatto, M. Entropic Dynamics: From entropy and information geometry to Hamiltonians and quantum mechanics. In Proceedings of the 34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Amboise, France, 21–26 September 2014; Mohammad-Djafari, A., Barbaresco, F., Eds.; AIP Publishing: Melville, NY, USA, 2015.
7. Johnson, D.T.; Caticha, A. Entropic dynamics and the quantum measurement problem. In Proceedings of the 31st International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Waterloo, ON, Canada, 10–15 July 2011; Knuth, K., Ed.; AIP Publishing: Melville, NY, USA, 2012; Volume 1443, pp. 104–116.
8. Johnson, D.T. Generalized Galilean Transformations and the Measurement Problem in the Entropic Dynamics Approach to Quantum Theory. Ph.D. Thesis, University at Albany, Albany, NY, USA, 26 April 2011.
9. Nawaz, S.; Caticha, A. Momentum and uncertainty relations in the entropic approach to quantum theory. In Proceedings of the 31st International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Waterloo, ON, Canada, 10–15 July 2011; Knuth, K., Ed.; AIP Publishing: Melville, NY, USA, 2012; Volume 1443, pp. 112–124.
10. Nawaz, S. Momentum and Spin in Entropic Quantum Dynamics. Ph.D. Thesis. University at Albany, Albany, NY, USA, 18 August 2014.

11. Ipek, S.; Caticha, A. Entropic Quantization of Scalar Fields. In Proceedings of the 34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Amboise, France, 21–26 September 2014; Mohammad-Djafari, A., Barbaresco, F., Eds.; AIP Publishing: Melville, NY, USA, 2015.

12. Bartolomeo, D.; Caticha, A. Entropic Dynamics: The Schrödinger equation and its Bohmian limit. In Proceedings of the 35th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Potstam, NY, USA, 19–24 July 2015.

13. Nawaz, S.; Abedi, M.; Caticha, A. Entropic Dynamics in Curved Spaces. In Proceedings of the 35th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Potstam, NY, USA, 19–24 July 2015.

14. Wootters, W.K. Statistical distance and Hilbert space. *Phys. Rev. D* **1981**, *23*, 357–362 .

15. Caticha, A. Consistency and Linearity in Quantum Theory. *Phys. Lett. A* **1998**, *244*, 13–17.

16. Caticha, A. Consistency, Amplitudes, and Probabilities in Quantum Theory. *Phys. Rev. A* **1998**, *57*, 1572–1582.

17. Caticha, A. Insufficient Reason and Entropy in Quantum Theory. *Found. Phys.* **2000**, *30*, 227–251.

18. Brukner, C.; Zeilinger, A. Information and Fundamental Elements of the Structure of Quantum Theory. In *Time, Quantum, Information*; Castell, L., Ischebeck, O., Eds.; Springer: Berlin, Germany, 2003.

19. Spekkens, R. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A* **2007**, *75*, 032110.

20. Goyal, P.; Knuth, K.; Skilling, J. Origin of complex quantum amplitudes and Feynman's rules. *Phys. Rev. A* **2010**, *81*, 022109.

21. Hardy, L. Reformulating and Reconstructing Quantum Theory. **2011**, arXiv.org:1104.2066.

22. Reginatto, M. From information to quanta: A derivation of the geometric formulation of quantum theory from information geometry. In Proceedings of the Symmetries in Science XVI, Bregenz, Austria, 21–26 July 2013.

23. Hall, M.J.W.; Reginatto, M. Schrödinger equation from an exact uncertainty principle. *J. Phys. A* **2002**, *35*, 3289–3299.

24. Hall, M.J.W.; Reginatto, M. Quantum mechanics from a Heisenberg-type inequality. *Fortschr. Phys.* **2002**, *50*, 646–656.

25. Nelson, E. *Quantum Fluctuations*; Princeton UP: Princeton, NJ, USA, 1985.

26. De la Peña, L.; Cetto, A.M. *The Quantum Dice: An Introduction to Stochastic Electrodynamics*; Kluwer: Dordrecht, The Netherland, 1996.

27. Smolin, L. Could quantum mechanics be an approximation to another theory? *High Energy Phys.* **2006**, arXiv.org/abs/quant-ph/0609109.

28. Hooft, G. Determinism beneath Quantum Mechanics. **2002**, arXiv:quant-ph/0212095.

29. Hooft, G. Emergent quantum mechanics and emergent symmetries. **2002**, arXiv:hep-th/0707.4568.

30. Elze, H.T.; Schipper, O. Time without time: A stochastic clock model. *Phys. Rev. D* **2002**, *66*, 044020.

31. Elze, H.T. Emergent discrete time and quantization: relativistic particle with extra-dimensions. *Phys. Lett. A* **2003**, *310*, 110–118.

32. Adler, S. *Quantum Theory as an Emergent Phenomenon*; Cambridge UP: Cambridge, UK, 2004.

33. Grössing, G. The Vacuum Fluctuation Theorem: Exact Schrödinger Equation via Nonequilibrium Thermodynamics. *Phys. Lett. A* **2008**, *372*, 4556–4563.

34. Grössing, G.; Fussy, S.; Pascasio, J.M.; Schwabl, H. The Quantum as an emergent System. *J. Phys. Conf. Ser.* **2012**, *361*, 012008.

35. Price, H. *Time's Arrow and Archimedes' Point*; Oxford UP: New York, NY, USA, 1996.

36. Zeh, H.D. *The Physical Basis of the Direction of Time*; Springer: Berlin, Germany, 2002.

37. Nelson, E. Review of Stochastic Mechanics. *J. Phys. Conf. Ser.* **2012**, *361*, 012011.

38. Nelson, E. Connection between Brownian motion and quantum mechanicsR. In *Einstein Symposium Berlin*; Lecture Notes Physis; Springer-Verlag: Berlin, Germany, 1979; Volume 100, p. 168.

39. Reginatto, M. Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information. *Phys. Rev. A* **1998**, *58*, 1775–1778.

40. Wallstrom, T.C. On the derivation of the Schrödinger equation from stochastic mechanics. *Found. Phys. Lett.* **1989**, *2*, 113–126.

41. Wallstrom, T.C. The inequivalence between the Schrödinger equation and the Madelung hydrodynamic equations. *Phys. Rev. A* **1994**, *49*, 1613–1616.

42. Takabayasi, T. Vortex, spin, and triad for quantum mechanics of spinning particle. *Prog. Theor. Phys.* **1983**, *70*, 1–17.