

Entropic Dynamics: an inference approach to quantum theory, time and measurement

Ariel Caticha

Department of Physics, University at Albany-SUNY, Albany, NY 12222, USA.

E-mail: ariel@albany.edu

Abstract. We review the derivation of quantum theory as an application of entropic methods of inference. The new contribution in this paper is a streamlined derivation of the Schrödinger equation based on a different choice of microstates and constraints.

1. An overview

Quantum mechanics involves probabilities in a fundamental way and, therefore, it is a theory of inference. But this has not always been clear. The controversy revolves around the interpretation of the quantum state — the wave function. Does it represent the actual real state of the system — its *ontic* state — or does it represent a state of knowledge about the system — an *epistemic* state?

Examples of ontic interpretations include, to name a few, Bohm’s causal interpretation, Everett’s many-worlds interpretation, and Nelson’s stochastic mechanics [1]. The epistemic interpretations have also had a number of advocates (for example, [2]-[5]) starting, most prominently, with Einstein. The “orthodox” or Copenhagen interpretation lies somewhere in between. On one hand, as described in standard textbooks such as the early classics by Dirac and von Neumann, it regards the quantum state as a complete and objective specification of the properties of the system — an ontic concept that is divorced from the state of knowledge of any rational agent. On the other hand, the founders of quantum theory — Bohr, Heisenberg, Born, etc. — were keenly aware of the epistemological and pragmatic elements in quantum mechanics (see e.g., [6]) but, unfortunately, they wrote at a time when the language, the tools and the rules of quantitative epistemology — the Bayesian and entropic methods of inference — had not yet been sufficiently developed. As a result they did not succeed in drawing a sharp line between the ontic and the epistemic and thereby started a controversy that lingers to this day.

But interpreting quantum theory is not merely a matter of postulating the mathematical formalism and then appending an interpretation to it. For the epistemic view of quantum states to be satisfactory it is not sufficient to state that the probability $|\psi|^2$ represents a state of knowledge; we must also show that changes or updates of the epistemic ψ — which include both the unitary time evolution according to the Schrödinger equation and the projection postulate during measurement — obey the rules of inference. In a truly epistemic interpretation there is no logical room for “quantum” probabilities obeying alternative rules of inference.

Our subject is Entropic Dynamics (ED), a framework in which quantum theory is formulated as an example of entropic inference [7]. ED differs from other approaches in several important



respects. For example, in the standard view quantum theory is considered as an extension of classical mechanics and therefore deviations from causality demand an explanation. In the entropic view, on the other hand, quantum mechanics is an example of entropic inference, a scheme designed to handle insufficient information [8]. From the entropic perspective indeterminism requires no explanation: uncertainty and probabilities are the norm. It is certainty and determinism that demand explanations.

ED also differs from other approaches based on information theory. (See *e.g.*, [9]-[15].) In ED the laws of physics are rules for processing information. The information in question possibly originates and might even find its ultimate justification in some sub-quantum dynamics that remains to be discovered. However, as we shall see, once the relevant information has been identified, the remaining details of any such underlying dynamics turn out to be irrelevant for behavior at the quantum level. In ED those irrelevant details are ignored from the start — which is a significant simplification. The situation is somewhat analogous to the laws of thermodynamics which also turns out to be largely independent of microscopic details at the atomic level.

The analogy with thermodynamics has inspired several attempts to explain the emergence of quantum behavior from specific proposals of a sub-quantum dynamics with some additional stochastic element. (See *e.g.*, [1][16]-[21].) In contrast, ED does not assume any underlying mechanics whether classical, deterministic, or stochastic. Both quantum dynamics and its classical limit are derived as examples of entropic inference.

Another difference is that ED naturally leads to an “entropic” notion of time. Time is introduced as a convenient book-keeping device to keep track of the accumulation of change. The task is to develop a model that includes (a) something one might identify as an “instant”, (b) a sense in which these instants can be “ordered”, (c) a convenient concept of “duration” measuring the separation between instants. The welcome new feature is that entropic time is intrinsically directional. Thus, an arrow of time is generated automatically.

ED offers a new perspective on the notorious problem of measurement (see [3][22][23]). Questions such as “How can a measurement ever yield a definite outcome?” or “Are the values of observables created during the act of measurement?” led von Neumann to postulate a dual mode of wave function evolution, either continuous and deterministic according to the Schrödinger equation, or discontinuous and stochastic during the measurement process. Once one accepts quantum theory as a theory of inference the dichotomy between the two modes disappears. Unitary evolution and discontinuous collapse correspond to two modes of updating probabilities which, as shown in [24], are not intrinsically different; they are special cases within a broader scheme of entropic inference [8].

Yet another distinguishing feature is that in ED the positions of particles have definite values just as they would in classical physics. This implies that the process of observation is essentially classical and measurements of position automatically yield definite outcomes. This solves the problem of measurement because position is the only observable. Indeed, in ED all other observables such as momentum, energy, and so on, are statistical concepts — just like temperature in statistical mechanics. They are not properties of the particles but of their probability distributions. As shown in [25][26] their values are indeed created by the act of measurement [27]. A more detailed treatment is given in [28] and [8].

In order to formulate quantum theory as an entropic dynamics — just as with any other inference problem — we must decide which microstates are the subject of our inference, we must identify the prior probabilities, and we must identify those constraints that represent the information that is relevant to our problem. The new contribution in this paper is an entropic derivation of the Schrödinger equation based on a choice of microstates and constraints that differs and is in some respects more advantageous than the choice adopted in [7].

2. Entropic Dynamics

In this model we consider particles living in flat three-dimensional space. The particles have *definite* positions and it is their unknown values that we wish to infer.¹

The basic dynamical assumptions are that motion happens and that it is continuous. Thus, short displacements happen and it is their accumulation that leads to motion. We do not explain why motion happens but, given the information that it does, our task is to venture a guess about what to expect.

For simplicity here we will focus on a single particle; the generalization to N particles is straightforward. For a single particle the configuration space is \mathbb{R}^3 with metric δ_{ab} . The particle moves from an initial x to an unknown x' . The goal is to find the probability distribution $P(x'|x)$. To find it maximize the appropriate (relative) entropy,

$$\mathcal{S}[P, Q] = - \int d^3x' P(x'|x) \log \frac{P(x'|x)}{Q(x'|x)}. \quad (1)$$

The relevant information is introduced through the prior probability $Q(x'|x)$, which reflects our knowledge about which x' to expect *before* we have any information about the motion, and the constraints that specify the family of acceptable posteriors $P(x'|x)$.

The prior We adopt a prior that represents a state of extreme ignorance. Knowledge of x tells us nothing about x' . Such ignorance is represented by a uniform distribution: $Q(x'|x)d^3x'$ is proportional to the volume element d^3x' . (The proportionality constant has no effect on the entropy maximization and can be safely ignored.)

The constraints on the motion The information that motion is continuous is imposed through a constraint. For a short step let $\vec{x}' = \vec{x} + \Delta\vec{x}$. We require that the expected squared displacement,

$$\langle \Delta\vec{x} \cdot \Delta\vec{x} \rangle = \kappa, \quad (2)$$

be some small but for now unspecified value κ , which we take to be independent of x in order to reflect the translational symmetry of the configuration space \mathbb{R}^3 .

If this were the only constraint the resulting motion would be a completely isotropic diffusion. Clearly some information is still missing. The additional piece of relevant information, that once particles are set in motion they tend to persist in it, is expressed by assuming the existence of a “potential” $\phi(x)$ and imposing that the expected displacement $\langle \Delta x^a \rangle$ in the direction of the gradient of ϕ is constrained to be

$$\langle \Delta\vec{x} \rangle \cdot \vec{\nabla} \phi = \kappa' \quad (3)$$

where κ' is another small but for now unspecified position-independent constant.

The seemingly ad hoc introduction of a potential ϕ will not be justified — at least not here. The important point is that we have identified the information needed for inference. Where this information originates and why it turns out to be relevant are, of course, interesting questions but their answers lie elsewhere — at some deeper level of physics. For the purpose of inference no further hypotheses need be made.²

¹ 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.

² In our earlier development of ED [7] the set of microstates involved the positions x and some additional mysterious variables that we called y . The present treatment is simpler in that no y variables need be postulated. This simplification comes at a price. In the y -variable model both the potential ϕ and the appearance of its gradient arise naturally without further assumptions. Thus we have a trade-off. We can simplify the microstates at the expense of the constraints. A deeper justification for ϕ , its geometric significance, and its natural relation to gauge symmetries can be given and will be discussed elsewhere in the context of particles with spin.

Having specified the prior and the constraints the ME method takes over. Varying $P(x'|x)$ to maximize $\mathcal{S}[P, Q]$ in (1) subject to the two constraints plus normalization gives

$$P(x'|x) = \frac{1}{\zeta} \exp\left[-\frac{1}{2}\alpha \Delta\vec{x} \cdot \Delta\vec{x} + \alpha' \Delta\vec{x} \cdot \vec{\nabla}\phi\right], \quad (4)$$

where ζ is a normalization constant,

$$\zeta(x, \alpha, \alpha') = \int d^3x' e^{-\frac{1}{2}\alpha \Delta\vec{x} \cdot \Delta\vec{x} + \alpha' \Delta\vec{x} \cdot \vec{\nabla}\phi}. \quad (5)$$

The Lagrange multipliers α and α' are determined in the standard way,

$$\partial \log \zeta / \partial \alpha = -\kappa/2 \quad \text{and} \quad \partial \log \zeta / \partial \alpha' = -\kappa'. \quad (6)$$

Since both the function ϕ and the constant κ' are so far unspecified, so is the multiplier α' . Without loss of generality, we can absorb α' into ϕ , $\alpha'\phi \rightarrow \phi$, which amounts to setting $\alpha' = 1$.

Eq.(4) for $P(x'|x)$ shows that short steps are obtained for large α and that they happen in essentially random directions with a small anisotropic bias along the gradient of ϕ . The distribution $P(x'|x)$ is Gaussian and is conveniently written as

$$P(x'|x) \propto \exp\left[-\frac{\alpha}{2} (\Delta\vec{x} - \langle \Delta\vec{x} \rangle)^2\right]. \quad (7)$$

The displacement $\Delta\vec{x} = \Delta\vec{x} + \Delta\vec{w}$ can be expressed as the expected drift plus a fluctuation

$$\langle \Delta\vec{x} \rangle = \Delta\vec{x} = \frac{1}{\alpha} \vec{\nabla}\phi, \quad (8)$$

$$\langle \Delta w^a \rangle = 0 \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{1}{\alpha} \delta^{ab}. \quad (9)$$

As $\alpha \rightarrow \infty$ the fluctuations become dominant: the drift $\Delta\vec{x} \sim \alpha^{-1}$ while $\Delta\vec{w} \sim \alpha^{-1/2}$. This implies that, as in Brownian motion, the trajectory is continuous but not differentiable. Here we see the roots of the uncertainty principle: a particle has a definite position but its velocity, the tangent to the trajectory, is completely undefined.

3. Entropic time

The foundation of all notions of time is dynamics. In ED time is introduced as a book-keeping device to keep track to the accumulation of small changes.

3.1. An ordered sequence of instants

In ED, at least for infinitesimally short steps, change is given by the transition probability $P(x'|x)$ in eq.(7). The n th step takes us from $x = x_{n-1}$ to $x' = x_n$. Using the product rule for the joint probability, $P(x_n, x_{n-1}) = P(x_n|x_{n-1})P(x_{n-1})$, and integrating over x_{n-1} , we get

$$P(x_n) = \int d^3x_{n-1} P(x_n|x_{n-1})P(x_{n-1}). \quad (10)$$

This equation is a direct consequence of the laws of probability. However, if $P(x_{n-1})$ happens to be the probability of different values of x_{n-1} at a given instant labelled t , then we will interpret $P(x_n)$ as the probability of values of x_n at the "later" instant $t' = t + \Delta t$. Accordingly, we write $P(x_{n-1}) = \rho(x, t)$ and $P(x_n) = \rho(x', t')$ so that

$$\rho(x', t') = \int d^3x P(x'|x)\rho(x, t) \quad (11)$$

Nothing in the laws of probability that led to eq.(10) forces this interpretation on us — this is an independent assumption about what constitutes time in our model. We use eq.(11) to define what we mean by an instant: *if the distribution $\rho(x,t)$ refers to one instant t , then the distribution $\rho(x',t')$ defines what we mean by the “next” instant $t' = t + \Delta t$* . Thus, eq.(11) allows *entropic time* to be constructed one instant after another.

We can phrase this idea somewhat differently. Once we have decided on the relevant information necessary for predicting future behavior we can imagine all that information codified into an “instant”. Thus, we define instants so that *given the present the future is independent of the past*.³

3.2. The arrow of entropic time

The notion of time as constructed according to eq.(11) is remarkable in that it incorporates an intrinsic directionality: there is an absolute sense in which $\rho(x,t)$ is prior and $\rho(x',t')$ is posterior.

Suppose we wanted to find a time-reversed evolution. We would write

$$\rho(x,t) = \int d^3x' P(x|x')\rho(x',t'). \quad (12)$$

This is perfectly legitimate but in order to be correct $P(x|x')$ cannot be obtained from eq.(7) by merely exchanging x and x' . According to the rules of probability theory $P(x|x')$ is related to eq.(7) by Bayes’ theorem,

$$P(x|x') = \frac{P(x)}{P(x')}P(x'|x). \quad (13)$$

In other words, one of the two transition probabilities, either $P(x'|x)$ or $P(x|x')$, *but not both*, can be given by the maximum entropy distribution eq.(7). The other is related to it by Bayes’ theorem. There is no symmetry between the inferential past and the inferential future because there is no symmetry between priors and posteriors.

The puzzle of the arrow of time has a long history (see *e.g.* [29][30]). The standard question has been how can an arrow of time be derived from underlying laws of nature that are symmetric? ED offers a new perspective. The asymmetry is the inevitable consequence of entropic inference. From the point of view of ED the challenge does not consist in explaining the arrow of time, but rather in explaining how it comes about that despite the arrow of time some laws of physics turn out to be reversible. Indeed, even when the derived laws of physics – in our case, the Schrödinger equation – turns out to be fully time-reversible, *entropic time itself only flows forward*.

3.3. Duration: a convenient time scale

Having introduced the notion of successive instants we now have to specify the interval Δt between them. This amounts to specifying the multiplier $\alpha(x,t)$ in terms of Δt .

Time is defined so that motion looks simple. For large α the dynamics is dominated by the fluctuations Δw . In order that the fluctuations $\langle \Delta w^a \Delta w^b \rangle$ reflect the symmetry of translations in space and time — a Newtonian time that flows “equally” everywhere and everywhen — we choose α to be independent of x and t , $\alpha(x,t) = C/\Delta t$, where C is some constant.

The extension of ED to several non-identical particles is not our subject here but a quick remark is useful. The extension is achieved by introducing separate constraints, eq.(2), for each particle, each with its own κ_i , and each with its own multiplier $\alpha_i = C_i/\Delta t$. It is convenient

³ An equation such as (11) is commonly employed to define Markovian behavior in which case it is known as the Chapman-Kolmogorov equation. Markovian processes are such that specifying the state of the system at time t is sufficient to fully determine its state after time t — no additional information about the past is needed. We make no Markovian assumptions. We are concerned with a different problem. We do not use (11) to define Markovian processes; we use it to define time.

to write each of these multipliers α_i as $\alpha_i = m_i/\hbar\Delta t$ in terms of a particle-specific constant m_i and an overall constant \hbar which fixes the units of the m_i s relative to the units of time. Thus

$$\alpha = \frac{m}{\hbar\Delta t} . \quad (14)$$

With this choice of the multiplier α the dynamics is indeed simple: $P(x'|x)$ in (7) becomes a standard Wiener process. The displacement is

$$\Delta\vec{x} = \vec{b}\Delta t + \Delta\vec{w} , \quad (15)$$

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

$$\langle\Delta\vec{x}\rangle = \vec{b}\Delta t \quad \text{with} \quad \vec{b} = \frac{\hbar}{m}\vec{\nabla}\phi , \quad (16)$$

and Δw^a is a fluctuation,

$$\langle\Delta w^a\rangle = 0 \quad \text{and} \quad \langle\Delta w^a\Delta w^b\rangle = \frac{\hbar}{m}\Delta t\delta^{ab} . \quad (17)$$

The formal similarity to Nelson's stochastic mechanics [1] is evident but the interpretations are completely different.

Two remarks are in order: one on the nature of clocks and another on the nature of mass.

On clocks: Time is defined so that motion looks simple. In Newtonian mechanics the prototype of a clock is the free particle. Time is defined so that the free particle moves 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 is defined so that the particle undergoes equal fluctuations in equal times.

On mass: The particle-specific constant m will, of course, be called 'mass' and eq.(17) provides its interpretation: mass is an inverse measure of fluctuations.

4. Accumulating changes: the Fokker-Planck equation

Equation is an integral equation for the evolution of the distribution $\rho(x, t)$. As is well-known from diffusion theory [8] it can be written in differential form as a Fokker-Planck equation (FP),

$$\partial_t\rho = -\vec{\nabla}\cdot(\rho\vec{b}) + \frac{\hbar}{2m}\nabla^2\rho , \quad (18)$$

which can itself be rewritten as a continuity equation,

$$\partial_t\rho = -\vec{\nabla}\cdot(\rho\vec{v}) . \quad (19)$$

The velocity \vec{v} of the probability flow or *current velocity* is

$$\vec{v} = \vec{b} + \vec{u} \quad \text{where} \quad \vec{u} = -\frac{\hbar}{m}\vec{\nabla}\log\rho^{1/2} , \quad (20)$$

the *osmotic velocity*, represents the tendency for probability to flow down the density gradient.

Since both \vec{b} and \vec{u} are gradients, it follows that the current velocity is a gradient too,

$$\vec{v} = \frac{\hbar}{m}\vec{\nabla}\Phi \quad \text{where} \quad \Phi = \phi - \log\rho^{1/2} . \quad (21)$$

With these results ED reaches a certain level of completion: We figured out what small changes to expect — they are given by $P(x'|x)$ — and time was introduced to keep track of how these small changes accumulate; the net result is diffusion according to the FP equation.

But quantum mechanics is not a *standard* diffusion. The discussion so far has led us to the density $\rho(x, t)$ as the important dynamical object but to construct a wave function, $\Psi = \rho^{1/2}e^{i\Phi}$, we need a second degree of freedom, the phase Φ . The problem is that as long as the potential ϕ is externally prescribed the function Φ in eq.(21) does not represent an independent degree of freedom. The natural solution is to relax this constraint and allow ϕ (or equivalently Φ) to participate in the dynamics. Thus the dynamics will consist of the coupled evolution of $\rho(x, t)$ and $\Phi(x, t)$.

5. Non-dissipative diffusion

To specify the dynamics we follow [31] and impose that the dynamics be non-dissipative, that is, we require the conservation of a certain functional $E[\rho, \phi]$ which will be called “energy”.

At first sight it might appear that imposing that some energy $E[\rho, \phi]$ be conserved is natural because it agrees with our classical preconceptions of what physics ought to be like. But classical intuitions are not a good guide here. In the more sophisticated approaches to physics energy is taken to be whatever happens to be conserved as a result of invariance under translations in time. But our dynamics has hardly been defined yet; what, then, is E and why should it be conserved in the first place? Furthermore, if we go back to eq.(15) we see that it is the kind of equation (a Langevin equation) that characterizes a Brownian motion in the limit of *infinite* friction. Thus, the explanation of quantum theory in terms of a sub-quantum classical mechanics would require that particles be subjected to infinite friction while suffering zero dissipation at the same time. Such a strange sub-quantum mechanics could hardly be called ‘classical’.

The energy functional $E[\rho, \phi]$ is chosen to be the expectation of a *local* “energy” function $\varepsilon(x, t)$, that is,

$$E[\rho, \phi] = \int d^3x \rho(x, t) \varepsilon(x, t) , \quad (22)$$

where $\varepsilon(x, t)$ depends on $\rho(x, t)$ and $\phi(x, t)$ and their derivatives.⁴ The local energy appropriate to the non-relativistic regime is

$$\varepsilon(x, t) = \frac{1}{2}mv^2 + \frac{1}{2}mu^2 + V(x) , \quad (23)$$

where the scalar function $V(x)$ represents an additional “potential” energy. The justification of ε is to be found in deeper-level physics but we can note that ε is tightly constrained by requiring invariance under time reversal ($\vec{v} \rightarrow -\vec{v}$ and $\vec{u} \rightarrow \vec{u}$) and the low velocity regime [32][7].

Using eqs.(20) and (21) the energy E is

$$E = \int d^3x \rho \left(\frac{\hbar^2}{2m} (\vec{\nabla}\Phi)^2 + \frac{\hbar^2}{2m} (\vec{\nabla} \log \rho^{1/2})^2 + V \right) \quad (24)$$

so that, after some algebra [7],

$$\frac{dE}{dt} = \int d^3x \dot{\rho} \left(\hbar\dot{\Phi} + \frac{\hbar^2}{2m} (\vec{\nabla}\Phi)^2 + V - \frac{\hbar^2}{2m} \frac{\nabla^2 \rho^{1/2}}{\rho^{1/2}} \right) \quad (25)$$

We impose that $\dot{E} = 0$ for spatially arbitrary choices of the initial conditions ρ and Φ , that is, at the initial t_0 we ought to be able to change ρ and Φ independently at different locations and still

⁴ In an energy eigenstate the local energy $\varepsilon(x, t)$ is uniform in space and constant in time.

get $\dot{E} = 0$. This implies the integrand should vanish at the initial t_0 . But any arbitrary time t can be taken as the initial time for evolution into the future. Therefore for all t we require that

$$\hbar\dot{\Phi} + \frac{\hbar^2}{2m}(\vec{\nabla}\Phi)^2 + V - \frac{\hbar^2}{2m}\frac{\nabla^2\rho^{1/2}}{\rho^{1/2}} = 0, \quad (26)$$

which is the quantum version of the Hamilton-Jacobi equation. Equations (26) and the FP equation,

$$\dot{\rho} = -\vec{\nabla} \cdot (\rho\vec{v}) = -\frac{\hbar}{m}\vec{\nabla} \cdot (\rho\vec{\nabla}\Phi) \quad (27)$$

are the coupled dynamical equations we seek.

These two real equations can be written as a single complex equation by combining ρ and Φ into a complex function $\Psi = \rho^{1/2} \exp(i\Phi)$. Computing the time derivative $\dot{\Psi}$ and using eqs.(26) and (27) leads to the Schrödinger equation,

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi. \quad (28)$$

Earlier we had introduced m as a particle-specific constant that measures fluctuations, \hbar as a constant that fixes units, and the entropic time t as a parameter designed to keep track of the accumulation of changes. Their relation to familiar physical quantities was a matter of conjecture. But now that we can see what role they play in the Schrödinger equation we can identify m with the particle mass, \hbar with Planck's constant, and the entropic time t with physical time.⁵

Other attempts to derive quantum theory start from an underlying, perhaps stochastic, classical mechanics. The ED approach is different in that it does not assume an underlying classical substrate; ED provides a derivation of *both Schrödinger's equation and also Newton's $F = ma$* . Classical mechanics is recovered in the usual limits of $\hbar \rightarrow 0$ or $m \rightarrow \infty$. Indeed, writing $S = \hbar\Phi$ in eq.(26) and letting $m \rightarrow \infty$ with S/m fixed leads to the classical Hamilton-Jacobi equation

$$\dot{S} + \frac{1}{2m}(\vec{\nabla}S)^2 + V = 0, \quad (29)$$

while eqs.(17), (20), and (21) give $m\vec{v} = \vec{\nabla}S$ and $\vec{u} = 0$ with vanishing fluctuations $\langle\Delta w^a\Delta w^b\rangle = \frac{\hbar}{m}\Delta t\delta^{ab} \rightarrow 0$.

6. Measurement in ED

In practice the measurement of position can be technically challenging because it requires the amplification of microscopic details to a macroscopically observable scale. However, no intrinsically quantum effects are involved: the position of a particle has a definite, albeit unknown, value x and its probability distribution is, by construction, given by the Born rule, $\rho(x) = |\Psi(x)|^2$. We can therefore assume that suitable position detectors are in principle available. First we consider observables other than position: how they are defined and how they are measured. Then we conclude with a few remarks on amplification and Bayes theorem.

6.1. Observables other than position

The fact that the Schrödinger equation (28) is linear and unitary makes the language of Hilbert spaces particularly convenient so from now we adopt Dirac's bra-ket notation and write $\Psi(x) = \langle x|\Psi\rangle$. For convenience we consider the case of a particle that lives on a discrete lattice.

⁵ Where by 'physical' we mean that it is the time t that appears in the laws of physics.

The generalization to a continuous space is straightforward. The probabilities of the previously continuous positions

$$\rho(x) d^3x = |\langle x|\Psi\rangle|^2 d^3x \quad \text{become} \quad p_i = |\langle x_i|\Psi\rangle|^2, \quad (30)$$

and if the state is

$$|\Psi\rangle = \sum_i c_i |x_i\rangle \quad \text{then} \quad p_i = |\langle x_i|\Psi\rangle|^2 = |c_i|^2. \quad (31)$$

Since position is the only objectively real quantity there is no reason to define other observables except that they turn out to be convenient when considering more complex experiments. Consider a setup in which right before reaching the position detector the particle is subjected to additional interactions, say magnetic fields or diffraction gratings. Suppose the interactions in such a complex setup \mathcal{A} are described by the Schrödinger eq.(28), that is, by a particular unitary evolution \hat{U}_A . The particle will be detected with certainty at position $|x_i\rangle$ provided it was initially in a state $|s_i\rangle$ such that

$$\hat{U}_A |s_i\rangle = |x_i\rangle. \quad (32)$$

Since the set $\{|x_i\rangle\}$ is orthonormal and complete, the corresponding set $\{|s_i\rangle\}$ is also orthonormal and complete,

$$\langle s_i|s_j\rangle = \delta_{ij} \quad \text{and} \quad \sum_i |s_i\rangle\langle s_i| = \hat{I}. \quad (33)$$

Now consider the effect of this setup \mathcal{A} on some generic initial state vector $|\Psi\rangle$ which can always be expanded as

$$|\Psi\rangle = \sum_i c_i |s_i\rangle, \quad (34)$$

where $c_i = \langle s_i|\Psi\rangle$ are complex coefficients. The state $|\Psi\rangle$ will evolve according to \hat{U}_A so that as it approaches the position detectors the new state is

$$\hat{U}_A |\Psi\rangle = \sum_i c_i \hat{U}_A |s_i\rangle = \sum_i c_i |x_i\rangle. \quad (35)$$

which, invoking the Born rule for position measurements, implies that the probability of finding the particle at the position x_i is

$$p_i = |c_i|^2 = |\langle s_i|\Psi\rangle|^2. \quad (36)$$

Thus, the probability that the particle in the initial state $|\Psi\rangle$ after going through the setup \mathcal{A} is found at position x_i is $|c_i|^2$.

The same experiment can be described from a point of view in which the setup \mathcal{A} is a black box, a complex detector the inner workings of which are not emphasized. The particle is detected at $|x_i\rangle$ *as if* it had earlier been in the state $|s_i\rangle$. We can adopt a new language and say, perhaps inappropriately, that the particle has effectively been “detected” in the state $|s_i\rangle$, and therefore, the probability that the particle in state $|\Psi\rangle$ is “detected” in state $|s_i\rangle$ is $|\langle s_i|\Psi\rangle|^2$ — which reproduces Born’s rule for a generic measurement device. The shift in language is not particularly fundamental — it is merely a matter of convenience but we can pursue it further and assert that the setup \mathcal{A} is a complex detector that “measures” all operators of the form

$$\hat{A} = \sum_i \lambda_i |s_i\rangle\langle s_i| \quad (37)$$

where the eigenvalues λ_i are arbitrary scalars.

Some remarks are in order. Note that when we say we have detected the particle at x_i *as if* it had earlier been in state $|s_i\rangle$ with eigenvalue λ_i we are not implying that the particle was in the particular state $|s_i\rangle$ — this is just a figure of speech. It is in this sense that the corresponding value λ_i of the observable \hat{A} has been “created by the act of measurement”. To be more explicit:

if a sentence such as “a particle has momentum \vec{p} ” is used only as a linguistic shortcut that conveys information about the wave function before the particle enters the complex detector then, strictly speaking, there is no such thing as the momentum of the particle. The momentum is not an attribute of the particle; it is an attribute of the epistemic state $\Psi(x)$.

Incidentally, note that it is not necessary that the eigenvalues of the operator \hat{A} be real — they could be complex numbers. What is necessary is that its eigenvectors $|s_i\rangle$ be orthogonal. This means that \hat{A} need not be Hermitian but its Hermitian and anti-Hermitian parts of \hat{A} must be simultaneously diagonalizable — they must commute.

In the standard interpretation of quantum mechanics Born’s rule (36) is a postulate; within ED it is the natural consequence of unitary time evolution and the fact that all measurements are ultimately position measurements. This raises the question of whether our scheme is sufficiently general to encompass all measurements of interest. While there is no general answer that will address all cases — who can, after all, even list all the measurements that future physicists might perform? — we can, nevertheless, ask whether our scheme includes a sufficiently large class of interesting measurements. How, for example, does one measure an observable for which there is no unitary transformation mapping its eigenstates to position eigenstates? Every case demands its own specific analysis. For example, how does one measure the energy of a free particle? A measurement device characterized by eigenvectors $\{|s\rangle\}$ measures all operators of the form $\hat{A} = \int ds \lambda(s)|s\rangle\langle s|$. Therefore the same device that measures the momentum \hat{p} of a particle (*e.g.*, using a magnetic field or a diffraction grating followed by a position detector such as a photographic plate or a photoelectric cell) can also be used to infer the energy $\hat{H} = \hat{p}^2/2m$ of a free particle.

Here is another example: It is not so easy to place a probe inside the atom, so how does one measure the energy of an electron that is bound to an atom? In practice the energy of the bound particle is not measured directly; instead it is inferred from the energy of photons emitted in transitions between the bound states. Since photons are free particles measuring their energy is not in principle problematic. This is a special case of the general scheme in which the system of interest and the pointer variable of an apparatus become correlated in such a way that observation of the pointer allows one to infer a quantity of the system. The paradigmatic example is a Stern-Gerlach experiment in which the particle’s position is the pointer variable that allows one to infer its spin.

The difficulty with the standard von Neumann interpretation is that it is not clear at what stage the pointer variable “collapses” and attains a definite value. This is precisely the difficulty of principle that is resolved in the entropic approach: the pointer variable is a position variable too and therefore always has a definite value.

7. Amplification

The technical problem of amplifying microscopic details so they can become macroscopically observable is usually handled with a detection device set up in an initial state of unstable equilibrium. The particle of interest activates the amplifying system by inducing a cascade reaction that leaves the amplifier in a definite macroscopic final state described by some pointer variable α .

An eigenstate $|s_i\rangle$ evolves to a position x_i and the goal of the amplification process is to infer the value x_i from the observed value α_r of the pointer variable. The design of the device is deemed successful when x_i and α_r are suitably correlated and this information is conveyed through a likelihood function $P(\alpha_r|x_i)$ — an ideal amplification device would be described by $P(\alpha_r|x_i) = \delta_{ri}$. Inferences about x_i follow from a standard application of Bayes’ rule,

$$P(x_i|\alpha_r) = P(x_i) \frac{P(\alpha_r|x_i)}{P(\alpha_r)} . \quad (38)$$

The point of these considerations is to emphasize that there is nothing intrinsically quantum mechanical about the amplification process. The issue is one of appropriate selection of the information (in this case the data α_r) that happens to be relevant to a certain inference (in this case x_i). This is, of course, a matter of design: a skilled experimentalist will design the device so that no spurious correlations—whether quantum or otherwise—nor any other kind of interfering noise will stand in the way of inferring x_i .

Acknowledgments

I would like to thank D. Bartolomeo, C. Cafaro, N. Caticha, S. DiFranzo, A. Giffin, P. Goyal, D. T. Johnson, K. Knuth, S. Nawaz, M. Reginatto, C. Rodríguez, and J. Skilling for many discussions on entropy, inference and quantum mechanics.

References

- [1] E. Nelson, *Quantum Fluctuations* (Princeton UP, Princeton 1985).
- [2] L. Ballentine, *Rev. Mod. Phys.* **42**, 358 (1970).
- [3] L. Ballentine, *Quantum Mechanics: A Modern Development* (World Scientific, Singapore 1998).
- [4] C. Caves, C. Fuchs, and R. Schack, *Studies in History and Philosophy of Modern Physics* **38**, 244 (2007).
- [5] N. Harrigan and R. Spekkens, *Found. Phys.* **40**, 125 (2010).
- [6] H. P. Stapp, *Am. J. Phys.* **40**, 1098 (1972).
- [7] A. Caticha, *J. Phys.* **A 44**, 225303 (2011) (arXiv.org/abs/1005.2357).
- [8] A. Caticha, *Entropic Inference and the Foundations of Physics* (USP Press, São Paulo, Brazil 2012); online at <http://www.albany.edu/physics/ACaticha-EIFP-book.pdf>.
- [9] W. K. Wootters, *Phys. Rev.* **D**, 357 (1981).
- [10] A. Caticha: *Phys. Lett.* **A244**, 13 (1998); *Phys. Rev.* **A57**, 1572 (1998); *Found. Phys.* **30**, 227 (2000).
- [11] C. Brukner and A. Zeilinger, “Information and Fundamental Elements of the Structure of Quantum Theory,” in *Time, Quantum, Information*, ed. L. Castell and O. Ischebeck (Springer, 2003) (arXiv:quant-ph/0212084).
- [12] R. Spekkens, *Phys. Rev.* **A 75**, 032110 (2007).
- [13] P. Goyal, K. Knuth, J. Skilling, *Phys. Rev.* **A 81**, 022109 (2010).
- [14] L. Hardy, “Reformulating and Reconstructing Quantum Theory” (arXiv.org:1104.2066).
- [15] M. Reginatto, “From information to quanta: a derivation of the geometric formulation of quantum theory from information geometry” (arxiv:1312.0429).
- [16] L. de la Peña and A. M. Cetto, *The Quantum Dice, an Introduction to Stochastic Electrodynamics* (Kluwer, Dordrecht, Holland 1996).
- [17] M. J. W. Hall and M. Reginatto, *J. Phys.* **A 35**, 3289 (2002); *Fortschr. Phys.* **50**, 646 (2002).
- [18] G. 't Hooft, “Determinism beneath Quantum Mechanics” (arxiv:quant-ph/0212095); “Emergent quantum mechanics and emergent symmetries” (arxiv:hep-th/0707.4568).
- [19] H. T. Elze and O. Schipper, *Phys. Rev.* **D66**, 044020 (2002) ; H. T. Elze, *Phys. Lett.* **A310**, 110 (2003).
- [20] S. Adler, *Quantum Theory as an Emergent Phenomenon* (Cambridge UP, Cambridge 2004).
- [21] G. Grössing, *Phys. Lett.* **A 372**, 4556 (2008) (arxiv:0711.4954); G. Grössing *et al*, “The Quantum as an emergent System” *J. Phys.: Conf. Ser.* **361** 012008 (2012).
- [22] M. Schlösshauer, *Rev. Mod. Phys.* **76**, 1267 (2004).
- [23] G. Jaeger, *Entanglement, Information, and the Interpretation of Quantum Mechanics* (Springer-Verlag, Berlin Heidelberg 2009).
- [24] A. Caticha and A. Giffin, “Updating Probabilities,” in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. by A. Mohammad-Djafari, *AIP Conf. Proc.* **872**, 31 (2006) (arXiv.org:physics/0608185).
- [25] A. Caticha, *Found. Phys.* **30**, 227 (2000).
- [26] D. T. Johnson and A. Caticha, “Entropic dynamics and the quantum measurement problem”, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. by K. Knuth *et al.*, *AIP Conf. Proc.* **1443**, 104 (2012) (arXiv:1108.2550).
- [27] S. Nawaz and A. Caticha, “Momentum and uncertainty relations in the entropic approach to quantum theory”, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. by K. Knuth *et al.*, *AIP Conf. Proc.* **1443**, 112 (2012) (arXiv:1108.2629).
- [28] D. T. Johnson, “Generalized Galilean Transformations and the Measurement Problem in the Entropic Dynamics Approach to Quantum Theory”, Ph.D. thesis, University at Albany (2011) (arXiv:1105.1384).
- [29] H. Price, *Time's Arrow and Archimedes' Point* (Oxford UP, New York 1996).

- [30] H. D. Zeh, *The Physical Basis of the Direction of Time* (Springer, 2002).
- [31] E. Nelson, "Connection between Brownian motion and quantum mechanics," in *Einstein Symposium Berlin*, Lect. Notes Phys. **100**, p.168 (Springer-Verlag, Berlin 1979).
- [32] L. Smolin, "Could quantum mechanics be an approximation to another theory?" ([arXiv.org/abs/quant-ph/0609109](https://arxiv.org/abs/quant-ph/0609109)).