A Simple Theory of Everything.

Shahid Nawaz*

October 3, 2024

Abstract

Using the method of maximum entropy we have obtained the following results — The exact solution of the Schrödinger equation is obtained for an arbitrary potential. The results are surprising. There are no particles. The entire universe is a compact dot that is probabilistic in one universe (dark matter and dark energy) while its wave function is zero in the dual universe (physical universe). Furthermore the fundamental constants of nature are constant in one universe while they are probabilities in the dual universe.

Keywords: Maximum entropy method, duality principle, renormalization, category theory, physical universe, dark matter and energy, and fundamental constants of nature.

1 Introduction

The paper is organized as follows. In section 2 we introduce the method of maximum entropy (MaxEnt) that was originally formulated by E. T. Jaynes [1, 2]. We then briefly discuss the generalization of MaxEnt in section 3. In section 4 we present Caticha's [3, 4] derivation of quantum mechanics using the framework of entropic dynamics (ED) while our actual work starts in the following sections. The paper is then concluded in section 11.

2 The Method of Maximum Entropy (MaxEnt)

In 1957, E. T. Jaynes [1, 2] formulated the method of maximum entropy (MaxEnt) to reconcile the statistical mechanics of J. W. Gibbs [5] and communication theory of C. E. Shannon [6].

The MaxEnt is a variational method that takes entropy is the starting point, and the goal is to find a candidate posterior that maximizes the entropy subject to certain constraints. To describe MaxEnt, here we follow Jaynes' original paper [1].

*snafridi@gmail.com

Consider a discrete case. The entropy of the probability distributions p_i is given by

$$S[p] = -\sum_i p_i \log p_i \,. \tag{1}$$

The probability distribution p_i is not given. All we know is the expectation value of certain functions f^k, for $k = 1, 2, \ldots$,

$$\langle f^k \rangle = \sum_i p_i f_i^k = F^k \,, \tag{2}$$

plus a normalization condition

$$\sum_i p_i = 1 \tag{3}$$

Next maximizing $S[p]$ subject to the constraints

$$0 = \delta\big(S[p] - \alpha \textstyle\sum_i p_i - \sum_k \lambda_k \sum_i p_i f_i^k\big) \,, \tag{4}$$

where λ's are Lagrange multipliers. The solution of eq. (4) is the generalized canonical distribution

$$p_i = \frac{\exp\left[-\sum_k \lambda_k f_i^k\right]}{Z} \,, \tag{5}$$

where Z is the partition function

$$Z = \sum_i \exp\left[-\textstyle\sum_k \lambda_k f_i^k\right] \tag{6}$$

For example, the Maxwell-Boltzmann distribution of statistical mechanics follows immediately if the only information available is the expected energy, $\langle E \rangle = \sum_i p_i E_i$, then

$$p_i = \frac{e^{-\beta E_i}}{Z} \,. \tag{7}$$

This concludes that statistical physics which is regarded as a physical theory is nothing but an example of inference.

3 Entropic Inference

The goal of inductive inference is to update from the prior to the posterior probability distribution when new information, either in the form of data or constraints, becomes available. Bayes' rule and MaxEnt are regarded as two parallel methods for update. Bayes' rule updates probability when the information is contained in arbitrary prior or in data, it cannot handle arbitrary constraints. On the other hand, MaxEnt can cope with arbitrary constraints but fixed prior. In MaxEnt the prior is nothing more than the underlying measure.

The MaxEnt method can be extended beyond its original scope. A full-fledged method for inductive inference is called the Maximum Entropy method (ME), also called entropic inference [7, 8] see also [3]. The entropic inference framework or ME is of general applicability. Whether the information is available in the form of data or constraints, it can be used for updating when new information becomes available. It turns out that both MaxEnt and Bayes' rule are special cases of ME.

In the entropic inference framework the probability distribution $p(x)$ should be ranked relative to the prior $q(x)$ according to the relative entropy

$$S[p,q] = -\int dx\, p(x)\, \log \frac{p(x)}{q(x)}\,. \tag{8}$$

4 Entropic Dynamics

In the Entropic Dynamics (ED) framework quantum theory is derived as an application of the method of maximum entropy [3, 4]. The goal is to do for quantum mechanics what Jaynes did for statistical mechanics.

ED is defined on the configuration space. It is assumed that the particles have definite positions x. For a single particle the configuration space X is Euclidean with the metric

$$\gamma_{ab} = \delta_{ab}/\sigma^2, \quad a,b = 1,2,3. \tag{9}$$

where σ^2 is a scale factor. The full significance of the scale factor only becomes apparent when discussing several particles with different masses [4].

In addition to the particle of interest there exists other variables which we call y and live in a space Y. We do not need to be very specific about the y variables, we will assume that their value is uncertain and that this uncertainty depends on the location x of the particle and is expressed by some probability distribution $p(y|x)$. We do not need to be very specific about $p(y|x)$ either. So we shall see it is their entropy that matters. The entropy of the y variables is given by

$$S[p,q] = -\int dy\, p(y|x) \log \frac{p(y|x)}{q(y)} = S(x)\,. \tag{10}$$

where $q(y)$ is some underlying measure which need not be specified further. Since x enters as a parameter in $p(y|x)$ the entropy is a function of x : $S[p,q] = S(x)$.

When the particle is allowed to move from an initial position x to a final position x', then the relevant space is $\mathcal{X} \times \mathcal{Y}$. In which case the appropriate entropy is

$$\mathcal{S}[P,Q] = -\int d^3x' dy' P(x',y'|x) \log \frac{P(x',y'|x)}{Q(x',y'|x)}\,, \tag{11}$$

where $Q(x',y'|x)$ is the prior probability distribution and $P(x',y'|x)$ is the posterior probability distribution. The acceptable posteriors can be obtained by making use of the prior information and specifying the relevant constraints.

The prior:
The prior probability distribution codifies relation between x' and y' given x before the actual information contained in the constraints has been processed. At this point we are ignorant about any relation between x' and y'. When the knowledge of x' tells us nothing about y' and vice versa, then the joint prior can be written as a product

$$Q(x', y'|x) = Q(x'|x)Q(y'|x) \,. \tag{12}$$

Furthermore we want to assign equal probabilities to equal volumes, that is,

$$Q(x'|x)d^3x' \propto \gamma^{1/2} d^3x' \,, \tag{13}$$

and

$$Q(y'|x)dy' \propto q(y')dy' \,. \tag{14}$$

Such distributions are called uniform distributions. Therefore up to a proportionality constant, the joint prior becomes

$$Q(x', y'|x) = \gamma^{1/2} q(y') \,, \tag{15}$$

where $\gamma = \det \gamma_{ab}$.
Constraints:
To specify the constraints, we write the posterior as

$$P(x', y'|x) = P(x'|x)P(y'|x', x) \tag{16}$$

The first constraint is introduced through the second factor in eq. (16) which codifies information about the uncertainty in y' given x, and x'. Since the particle does not remember the past history, the uncertainty in y' must only depend on the later position x'. This means that

$$P(y'|x', x) = p(y'|x') \,, \tag{17}$$

where $p(y'|x')$ is the probability distribution of y variables.

The second constraint concerns the factor $P(x'|x)$ in eq. (16) which represents the transition probability from x to x'. We require that actual physical changes happen continuously, there is no discontinuity while moving from x to x'. To allow the continuity condition we require that x' is infinitesimally close to x. This information is incorporated in to the following constraint: Let $\Delta x = x' - x$, then the expectation

$$\left\langle \Delta\ell^2 \right\rangle = \left\langle \gamma_{ab} \Delta x^a \Delta x^b \right\rangle \,, \tag{18}$$

be some small numerical value, which we take to be independent of x in order to reflect the translational symmetry of the space $\mathcal{X}$.

The last constraint involves the normalization condition

$$\int d^3x' \, P(x'|x) = 1 \tag{19}$$

Taking into account the prior (15) and the constraint (17), the joint entropy (11) takes the form

$$\mathcal{S}[P,Q] = -\int d^3x' P(x'|x) \log \frac{P(x'|x)}{\gamma^{1/2}} + \int dx' P(x'|x) S(x') \,, \tag{20}$$

where $S(x)$ is given by eq. (10).

Next we vary $P(x'|x)$ to maximize $\mathcal{S}[P,Q]$ subject to the additional constraints (18) and (19). The result is

$$P(x'|x) = \frac{1}{\zeta} e^{S(x') - \frac{1}{2}\alpha(x)\Delta\ell^2} \,, \tag{21}$$

where ζ is a normalization constant and α is a Lagrange multiplier.

The transition probability $P(x'|x)$ is meant to hold for short steps. This happens when α is very large. For large α, eq. (21) can be approximated to

$$P\left(x'|x\right) \approx \frac{1}{Z} \exp\left[-\frac{\alpha(x)}{2\sigma^2}\delta_{ab}\left(\Delta x^a - \Delta\bar{x}^a\right)\left(\Delta x^b - \Delta\bar{x}^b\right)\right] . \tag{22}$$

where Z is a new normalization constant. The displacement Δx^a can be expressed as an expected drift plus a fluctuation,

$$\Delta x^a = \Delta\bar{x}^a + \Delta w^a \,, \tag{23}$$

where

$$\langle \Delta x^a \rangle = \Delta\bar{x}^a = \frac{\sigma^2}{\alpha(x)} \delta^{ab} \partial_b S(x) \,, \tag{24}$$

$$\langle \Delta w^a \rangle = 0 \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{\sigma^2}{\alpha(x)} \delta^{ab} \,. \tag{25}$$

As can be seen from eq. (24), the particle tends to drift along the entropy gradient. For large α the step size becomes very small but the fluctuations become dominant, because $\Delta\bar{x} \sim O(\alpha^{-1})$ while $\Delta w \sim O(\alpha^{-1/2})$. It means that as $\alpha \to \infty$ the trajectory is continuous but not differentiable—just like Brownian motion.

The Construction of Entropic Time:

The concept of time is closely connected with motion and change [9]. In entropic dynamics (ED) motion is described by the transition probability, eq. (22), that takes in to account small changes in short steps. On the other hand, larger changes are the accumulation of very many small short steps.

To construct time in ED we note that any notion of time has several aspects: (a) an instant of time, (b) the temporal order of instants, (c) the duration of time [10]. We begin with constructing an instant of time. Consider the particle is initially at position x and it moves to a final position x'. In general both x and x' are unknown. This means that we must deal with the joint probability $P(x, x')$, and then using the product rule

$$P(x', x) = P(x'|x) P(x) \,. \tag{26}$$

We note that $P(x'|x)$ is the probability of x' given x, but x is also unknown so we marginalize over x

$$P(x') = \int P(x', x)dx = \int P(x'|x)P(x)dx\,, \tag{27}$$

where $P(x)$ is the probability of the particle being located at position of x and $P(x')$ is the probability of the particle being found at x'. Since x is the initial position which occurs at an initial time t and x' occurs at a later time $t' > t$, therefore we write $P(x) = \rho(x, t)$ and $P(x') = \rho(x', t')$ so that

$$\rho(x', t') = \int P(x'|x)\rho(x, t)dx\,, \tag{28}$$

where t and t' are different instants of time which are ordered according earlier and later ($t' > t$).

Having introduced the notion of time in entropic dynamics the next important issue is of the duration or interval of time. Since we want to reconstruct non relativistic quantum mechanics, we need to construct Newtonian time. In Newtonian time, time flows equably independent of position x.. To achieve this we assume that the Lagrange multiplier α is a constant such that

$$\alpha = \frac{\tau}{\Delta t} = \text{constant}\,, \tag{29}$$

where τ is a constant that sets the unit of time interval Δt.

Finally the transition probability, eq. (22), becomes

$$P\left(x'|x\right) \approx \frac{1}{Z} \exp\left[-\frac{\tau}{2\sigma^2 \Delta t}\delta_{ab}\left(\Delta x^a - \Delta\bar{x}^a\right)\left(\Delta x^b - \Delta\bar{x}^b\right)\right]\,. \tag{30}$$

which can be recognized as standard Wiener process where now eq. (23) can be expresses in a familiar form

$$\Delta x^a = b^a\left(x\right)\Delta t + \Delta w^a, \tag{31}$$

where

$$b^a\left(x\right) = \frac{\sigma^2}{\tau}\delta^{ab}\partial_b S\left(x\right)\,, \tag{32}$$

is the drift velocity, and

$$\langle\Delta w^a\rangle = 0 \quad \text{and} \quad \left\langle\Delta w^a \Delta w^b\right\rangle = \frac{\sigma^2}{\tau}\Delta t \delta^{ab}\,, \tag{33}$$

where $\sigma^2/2\tau$ is the diffusion constant.

Derivation of the Schrödinger Equation:

The set of equations (30-33) describe small changes. Standard methods show that the successive iteration of eq. (28) yields a probability distribution $\rho\left(x, t\right)$ that evolves according to Fokker-Planck equation [11, 12, 3]

$$\frac{\partial\rho}{\partial t} = -\partial_a\left(b^a\rho\right) + \frac{\sigma^2}{2\tau}\nabla^2\rho\,, \tag{34}$$

which can be written as an equation for conservation of probability

$$\partial_t \rho = -\partial_a \left(\rho v^a \right) \tag{35}$$

Clearly v^a is interpreted as the velocity of flow of probability — it is called the current velocity. The current velocity can also be written as

$$v^a = b^a - \frac{\sigma^2}{2\tau} \delta^{ab} \frac{\partial_b \rho}{\rho} \,, \tag{36}$$

where b^a is the drift velocity given by eq. (32). The drift velocity reflects flow up the entropy gradient.

The second term in eq. (36) can be conveniently defined as

$$u^a = -\frac{\sigma^2}{\tau} \delta^{ab} \partial_b \log \rho^{1/2} \,, \tag{37}$$

To interpret eq. (37) we write it as

$$\rho u^a = -\frac{\sigma^2}{2\tau} \delta^{ab} \partial_b \rho \,, \tag{38}$$

which we recognize as Fick's Law and shows that ρu^a is the probability flux due to diffusion. The velocity u^a is called the osmotic velocity.

The current velocity can also be written as

$$v^a = \frac{\sigma^2}{\tau} \delta^{ab} \partial_b \phi \ , \quad \text{with} \quad \phi \left(x,t \right) = S \left(x \right) - \log \rho^{1/2} \left(x,t \right) \tag{39}$$

which shows that the current velocity is a gradient. In Nelson theory the current velocity was postulated to be a gradient. In ED, this fact is derived!

The dynamics described so far does not fully describe a diffusion. We note that the kinetic energy $\frac{1}{2} m \left(\frac{dx}{dt} \right)^2$ is infinite because dx/dt is infinite. This means that the energy is not conserved. Further assumption is needed to overcome this problem. Here we borrow Nelson's brilliant idea that diffusion can be non-dissipative if the expected energy is conserved [13]. In entropic dynamics, this constraint means allowing $p(y|x)$ and $S\left(x\right)$ to be functions of time, $S = S\left(x,t\right)$.

To this end introduce an energy functional [4, 3],

$$E[\rho, S] = \int d^3 x \rho \left(x,t \right) \left(\frac{1}{2} m v^2 + \frac{1}{2} \mu u^2 + V \left(x \right) \right) \,, \tag{40}$$

where m and μ are constants that will be called the mass and the osmotic mass respectively.

When the potential is static $\dot{V} = 0$, then for the arbitrary initial choices of ρ and ϕ the energy conservation ($\dot{E} = 0$) leads to the quantum Hamilton-Jacobi equation,

$$\eta \dot{\phi} + \frac{\eta^2}{2m} \left(\partial_a \phi \right)^2 + V - \frac{\mu \eta^2}{2m^2} \frac{\nabla^2 \rho^{1/2}}{\rho^{1/2}} = 0 \,, \tag{41}$$

where we have defined a new constant η so that

$$\eta \stackrel{\text{def}}{=} m\sigma^2/\tau\,, \tag{42}$$

In terms of η, the Focker-Planck equation (35) becomes

$$\dot{\rho} = -\frac{\eta}{m}\partial^a(\rho\partial_a\phi)\,. \tag{43}$$

Eqs. (41) and (43) are the entropic dynamical equations that determine the evolution of the dynamical variables $\phi(x,t)$ and $\rho(x,t)$.

It should be noted that eq. (41) can be obtained without loss of generality even when the potential is time dependent in which case the energy increases at the rate of

$$\dot{E} = \int d^3x\rho\dot{V}\,. \tag{44}$$

The two couple equations (41) and (43), which involve real quantities, can be combined into a single complex equation by introducing a complex quantity

$$\Psi = \rho^{1/2}e^{i\phi}\,, \tag{45}$$

then

$$i\eta\dot{\Psi} = -\frac{\eta^2}{2m}\nabla^2\Psi + V\Psi + \frac{\eta^2}{2m}\left(1-\frac{\mu}{m}\right)\frac{\nabla^2\left(\Psi\Psi^*\right)^{1/2}}{\left(\Psi\Psi^*\right)^{1/2}}\Psi\,. \tag{46}$$

This reproduces Schrödinger equation provided $\mu = m$,

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi\,, \tag{47}$$

where we have also identified η with $\hbar$.

There are essentially two possibilities if $m \neq \mu$: either $\mu > 0$ or $\mu = 0$. Here we analyze both cases separately. First we consider the former case. It turns out that all theories with $\mu > 0$ are physically equivalent in that they can be regraduated to a theory with $\mu_{\text{new}} = m$. To show this we note that the units η and τ can always be rescaled into $\eta = \kappa\eta'$ and $\tau = \kappa\tau'$ while simultaneously rescaling ϕ into $\phi = \phi'/\kappa$ where κ is some constant. Making these substitutions in eqs. (43) and (40) we get

$$\frac{\partial\rho}{\partial t} = -\frac{\eta'}{m}\partial_a\left(\rho\partial^a\phi'\right)\,, \tag{48}$$

and

$$E[\rho,S] = \int d^3x\rho\left(\frac{\eta'^2}{2m}\left(\partial_a\phi'\right)^2 + \frac{\mu\kappa^2\eta'^2}{8m^2}\left(\partial_a\log\rho\right)^2 + V\right)\,. \tag{49}$$

Again follow the same procedure that led to eq. (46) we get

$$i\eta'\dot{\Psi}' = -\frac{\eta'^2}{2m}\nabla^2\Psi' + V\Psi + \frac{\eta'^2}{2m}\left(1-\frac{\mu\kappa^2}{m}\right)\frac{\nabla^2\left(\Psi'\Psi'^*\right)^{1/2}}{\left(\Psi'\Psi'^*\right)^{1/2}}\Psi'\,, \tag{50}$$

where now $\Psi' = \rho^{1/2} e^{i\phi'}$. Since κ is just a rescaling factor which has no physical implications we can tune it so that $\mu_{\text{new}} = \mu\kappa^2 = m$, and thus we again recover the Schrödinger equation provided $\mu = m$,

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V\Psi \,, \tag{51}$$

where we dropped primes over Ψ and identifying η' with $\hbar$.

The other possibility occurs for $\mu = 0$ which allows no regraduation and leads to a non-linear Schrödinger equation,

$$i\hbar \dot{\Psi} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V\Psi + \frac{\hbar^2}{2m} \frac{\nabla^2 \left(\Psi\Psi^*\right)^{1/2}}{\left(\Psi\Psi^*\right)^{1/2}} \Psi \,. \tag{52}$$

This case is explored in [14] that exhibits both classical and quantum features.
External Electromagnetic Field:
Entropic dynamics can handle an external electromagnetic field in a natural way. If the particle is placed in an external field, it constrains the possible trajectories of the particle. To encode this additional information in the transition probability, the following constraint is to be used

$$\langle \Delta x^a A_a(x) \rangle = C \,, \tag{53}$$

where $A_a(x)$ are the components of the vector potential and C is a constant. This constraint only allows the expected components of displacements along the direction of $A_a(x)$.

Carrying out the calculations as in the previous sections, the transition probability turns out to be [4, 3],

$$P(x'|x) \propto \exp\left[-\frac{m}{2\hbar\Delta t} \delta_{ab} \left(\Delta x^a - \Delta\bar{x}^a\right) \left(\Delta x^b - \Delta\bar{x}^b\right) \right] \,, \tag{54}$$

where the displacement Δx^a is given by

$$\Delta x^a = \Delta\bar{x}^a + \Delta w^a \,, \tag{55}$$

with

$$\Delta\bar{x}^a = b^a \Delta t \quad \text{where} \quad b^a = \frac{\hbar}{m} \delta^{ab} [\partial_b S - \lambda A_b] \,, \tag{56}$$

where λ is a Lagrange multiplier that arises due to the additional constraint, eq. (53). The fluctuations are given by

$$\langle \Delta w^a \rangle = 0 \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{\hbar}{m} \Delta t \delta^{ab} \,. \tag{57}$$

The Fokker-Planck equation takes the form

$$\dot{\rho} = -\partial_a (\rho v^a) \,, \tag{58}$$

where now the current velocity is given by

$$v^a = \frac{\hbar}{m}\delta^{ab}(\partial_b\phi - \lambda A_b)\,. \tag{59}$$

While the forms of ϕ and the osmotic velocity u^a do not change, that is

$$\phi(x,t) = S(x,t) - \log\rho^{1/2}(x,t)\,, \tag{60}$$

and

$$u^a = -\frac{\hbar}{m}\delta^{ab}\partial_b \log\rho^{1/2}\,. \tag{61}$$

The energy functional is the same as in eq. (40), but now the current velocity is given by eq. (59),

$$E = \int d^3x\rho\left(\frac{\hbar^2}{2m}(\partial_a\phi - \lambda A_a)^2 + \frac{\hbar^2}{8m}(\partial_a \log\rho)^2 + V\right)\,. \tag{62}$$

When the external potentials are static, $\dot{V} = 0$ and $\dot{A} = 0$, then the energy conservation $\dot{E} = 0$ leads to the following equation

$$\hbar\dot{\phi} + \frac{\hbar^2}{2m}(\partial_a - \lambda A_a)^2 + V - \frac{\hbar^2}{2m}\frac{\nabla^2\rho^{1/2}}{\rho^{1/2}} = 0\,. \tag{63}$$

It should be noted again that eq. (63) can obtained without loss of generality. When the external potentials are time dependent then require that the energy increase at the rate

$$\dot{E} = \int d^3x\rho(\dot{V} + \hbar\lambda v^a\dot{A}_a) \tag{64}$$

Now again let $\Psi = \rho^{1/2}e^{i\phi}$, then eqs. (58) and (63) lead to the Schrödinger equation in an external electromagnetic field,

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{\hbar^2}{2m}(i\partial_a - \lambda A_a)^2\Psi + V\Psi \tag{65}$$

It turns out that the Lagrange multiplier λ plays the rule of electric charge e by making the identification $\lambda = e/\hbar c$.

5 Extended Entropic Dynamics (EED)

Here we present our main work.

We again consider a single particle that lives in a configuration space X. We do not require that X be flat. Let $\gamma_{ab}(x)$ be the metric defined on space $\mathcal{X}$. Next we allow that changes indeed happen that is the particle moves from an initial position x to an unknown position x'. The future position x' is known, however, it is somewhere between x and $x + \Delta x$. Therefore x' belongs to an interval

$$x' \in I_{x'} = \{x' : a < x' < b\}\,. \tag{66}$$

It means that there are infinite possible future points. For each x there is a corresponding y in y variables

$$y' \in I_{y'} = \{y' : c < y' < d\}\,, \tag{67}$$

and for each x there is a corresponding $p(y|x)$. The entropy $S(x)$ of $p(y|x)$ relative to an underlying measure $q(y)$ of the space $\mathcal{Y}$ is

$$S(x) = -\int dy\, p(y|x) \log \frac{p(y|x)}{q(y)}\,, \tag{68}$$

which is the same as in eq. (10). Since there are infinite possible future positions, we want to find the joint distribution $P(I_{x'}, I_{y'}|x)$ and the appropriate entropy is

$$\mathcal{S}[P,Q] = -\int Dx' Dy' P(I_{x'}, I_{y'}|x) \log \frac{P(I_{x'}, I_{y'}|x)}{Q(I_{x'}, I_{y'}|x)}\,. \tag{69}$$

Two points case:
Let us start with the possibility that the particle can only move to two possible future positions. The result will be generalized for the whole interval $I_{x'}$ at the end. Here we assume that the particle can either move to $x' \in I_{x'}$ or $x'' \in I_{x'}$ so that

$$\mathcal{S}[P,Q] = -\int dx'dy'dx''dy'' P(x', y', x'', y''))|x) \log \frac{P(x', y', x'', y''|x)}{Q(x', y', x'', y''|x)}\,. \tag{70}$$

The prior:
We select a prior that represents a state of extreme ignorance: the relations between x' and y', and x'' and y'' are not known. Such ignorance is represented by a product

$$Q(x', y', x'', y''|x) = Q(x'|x)Q(y'|x)Q(x''|x)Q(y''|x)\,. \tag{71}$$

Furthermore we take the distributions $Q(x'|x)d^3x'$, $Q(y'|x)dy'$ etc. to be uniform. Therefore, up to an irrelevant proportionality constant, the joint prior is

$$Q(x', y', x'', y''|x) = \gamma^{1/2}(x')q(y')\gamma^{1/2}(x'')q(y'')\,. \tag{72}$$

The constraints:
Before we specify the constraints we write the joint posterior as

$$P(x', y', x'', y''|x) = P(x'|x)P(y'|x', x)P(x''|x', y', x)P(y''|x'', x', y', x)\,. \tag{73}$$

The First constraint:
The first constraint deals with with second factor in eq. (73). We require that x' and y' be related in very specific way, namely that

$$P(y'|x', x) = p(y'|x')\,, \tag{74}$$

where $p(y'|x')$ is the probability of y variables. This is the same constraint used in section 4, namely eq. (17).

The second constraint:
The second constraint concerns the third factor in eq. (73). We require that x'' is unrelated to y',

$$P(x''|x',y',x) = P(x''|x',x)\,. \tag{75}$$

The third constraint:
The third constraint concerns the last factor in eq. (73). We require that y'' is only related to x'' so that

$$P(y''|x'',x',y',x) = p(y''|x'')\,. \tag{76}$$

where $p(y''|x'')$ is the probability of of y variables. This constraint is similar to constraint (74).

Substituting the prior (72) and the constraints (74 – 76) in the joint entropy (70). The result is

$$\mathcal{S}[P,Q] = T_1 + T_2 + T_3 + T_4\,, \tag{77}$$

where

$$T_1 = -\int dx' P(x'|x)\log\frac{P(x'|x)}{\gamma^{1/2}(x')}\,, \tag{78}$$

$$T_2 = \int dx' P(x'|x)S(x')\,, \tag{79}$$

where we have also used eq. (68).

$$T_3 = -\int dx'dx'' P(x'|x)P(x''|x',x)\log\frac{P(x''|x',x)}{\gamma^{1/2}(x'')}\,. \tag{80}$$

Similarly

$$T_4 = \int dx'dx'' P(x'|x)P(x''|x',x)S(x'')\,. \tag{81}$$

Consider eq. (80). We want to expand $P(x''|x',x)$ about x. However if

$$x'' = x + \Delta x\,, \tag{82}$$

then

$$P(x''|x',x) \to \delta(x''-x) \quad \text{as} \quad \Delta x \to 0\,. \tag{83}$$

It involves singular behavior. The singular behavior can be avoided if one instead takes

$$x' = x + \Delta x\,, \tag{84}$$

and then expand $P(x''|x',x)$ about x. It involves the distinguishability of two neighboring distributions $P(x''|x)$ and $P(x''|x+\Delta x)$. Therefore we assume eq. (84), that is $\Delta x = x' - x$, and then expand $P(x''|x',x)$ in eq. (80) about x,

$$T_3 = -\int dx'dx'' P(x'|x)P(x''|x)F\,, \tag{85}$$

where

$$F = (1 + \Delta x^a \partial_a \log P(x''|x) + \cdots) \log \frac{P(x''|x)\,(1 + \Delta x^a \partial_a \log P(x''|x))}{\gamma^{1/2}(x'')} \,. \quad (86)$$

Write

$$\begin{aligned} \log \frac{P(x''|x)\,(1 + \Delta x^a \partial_a \log P(x''|x))}{\gamma^{1/2}(x'')} &= \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} \\ &\quad + \log(1 + \Delta x^a \partial_a \log P(x''|x)) \,. \quad (87) \end{aligned}$$

Since

$$\log(1+x) = x - \frac{x^2}{2} + \cdots \,, \quad (88)$$

therefore

$$\begin{aligned} \log \frac{P(x''|x)\,(1 + \Delta x^a \partial_a \log P(x''|x))}{\gamma^{1/2}(x'')} &= \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} \\ &\quad + \Delta x^a \partial_a \log P(x''|x) \\ &\quad - \frac{1}{2} \partial_a \log P(x''|x) \partial_b \log P(x''|x) \Delta x^a \Delta x^b \,. \quad (89) \end{aligned}$$

and

$$\begin{aligned} F &= \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} + \partial_a \log P(x''|x) \Delta x^a \\ &\quad + \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} \partial_a \log P(x''|x) \Delta x^a \\ &\quad + \frac{1}{2} \partial_a \log P(x''|x) \partial_b \log P(x''|x) \Delta x^a \Delta x^b \,. \quad (90) \end{aligned}$$

Substituting eq. (103) in eq. (85), we have

$$\begin{aligned} T_3 &= -\int dx' dx'' P(x'|x) P(x''|x) \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} \\ &\quad - \int dx' dx'' P(x'|x) P(x''|x) \partial_a \log P(x''|x|x) \Delta x^a \\ &\quad - \int dx' dx'' P(x'|x) P(x''|x) \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} \partial_a \log P(x''|x) \Delta x^a \\ &\quad - \frac{1}{2} \int dx' dx'' P(x'|x) P(x''|x) \partial_a \log P(x''|x) \partial_b \log P(x''|x) \Delta x^a \Delta x^b \,. \quad (91) \end{aligned}$$

Consider the first term on the r. h. s. of eq. (91)

$$-\int dx' dx'' P(x'|x) P(x''|x) \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} = -\int dx'' P(x''|x) \log \frac{P(x''|x)}{\gamma^{1/2}(x'')} \,, \quad (92)$$

where we have performed the x' integration. Furthermore, since x'' on the r. h. s. of eq. (92) is a dummy variable, we can redefine it as x' so that

$$-\int dx'dx''P(x'|x)P(x''|x)\log\frac{P(x''|x)}{\gamma^{1/2}(x'')} = -\int dx'P(x'|x)\log\frac{P(x'|x)}{\gamma^{1/2}(x')} \,. \quad (93)$$

The second term on the r. h. s. of eq. (91) vanishes because

$$\int dx''P(x''|x)\partial_a \log P(x''|x|x)\Delta x^a = \Delta x^a \frac{\partial}{\partial x^a}\int dx''P(x''|x) = 0\,. \quad (94)$$

In the third on the r. h. s. of eq. (91), define

$$A_a(x) \overset{\text{def}}{=} \int dx''P(x''|x)\log\frac{P(x''|x)}{\gamma^{1/2}(x'')}\partial_a \log P(x''|x)\,, \quad (95)$$

so that

$$-\int dx'dx''P(x'|x)P(x''|x)\log\frac{P(x''|x)}{\gamma^{1/2}(x'')}\partial_a \log P(x''|x)\Delta x^a = -\int dx'P(x'|x)A_a(x)\Delta x^a\,. \quad (96)$$

The quantity A_a will be interpreted later.

The forth term on the r. h. s. of eq. (91) involves information metric

$$g_{ab}(x) = \int dx''P(x''|x)\partial_a \log P(x''|x)\partial_b \log P(x''|x)\,. \quad (97)$$

The metric $g_{ab}(x)$ describes the distance $d\ell$ between two neighboring distributions $P(x''|x)$ and $P(x''|x+\Delta x)$ or, equivalently $g_{ab}(x)$ also describes the same distance $d\ell$ between two points x and $x+dx$ [3]. Since x and $x+dx$ are points in space $\mathcal{X}$, while previously the metric of space $\mathcal{X}$ was given by $\gamma_{ab}(x)$. Therefore $g_{ab}(x)$ is the same as $\gamma_{ab}(x)$,

$$g_{ab}(x) = \gamma_{ab}(x) \quad (98)$$

Finally eq. (91) becomes

$$\begin{aligned} T_3 \;=\; & -\int dx'P(x'|x)\log\frac{P(x'|x)}{\gamma^{1/2}(x')} \\ & -\int dx'P(x'|x)A_a(x)\Delta x^a \\ & -\frac{1}{2}\int dx'P(x'|x)\gamma_{ab}(x)\Delta x^a\Delta x^a\,. \end{aligned} \quad (99)$$

Now recall eq. (81)

$$T_4 = \int dx'dx''P(x'|x)P(x''|x',x)S(x'')\,. \quad (81)$$

Now again assume eq. (84), expand $P(x''|x',x)$ about x so that

$$P(x''|x',x) = P(x''|x)\left(1 + \Delta x^a \partial_a \log P(x''|x) + \cdots\right) , \tag{100}$$

so that

$$\begin{aligned} T_4 &= \int dx' P(x'|x) \int dx'' P(x''|x) S(x'') \\ &\quad + \int dx' P(x'|x) \Delta x^a \frac{\partial}{\partial x^a} \int dx'' P(x''|x) S(x'') \,. \end{aligned} \tag{101}$$

In the first term perform the integration over x' and then redefine x'' as x' so that

$$\int dx' P(x'|x) \int dx'' P(x''|x) S(x'') = \int dx' P(x'|x) S(x') \,. \tag{102}$$

In the second term of eq. (101), define

$$F_a(x) \stackrel{\text{def}}{=} \frac{\partial}{\partial x^a} \int dx'' P(x''|x) S(x'') \,. \tag{103}$$

Finally

$$T_4 = \int dx' P(x'|x) S(x') + \int dx' P(x'|x) F_a(x) \Delta x^a(x), . \tag{104}$$

Collecting the results. Substitute eqs. (78), (79), (99) and (104) in the joint entropy (77),

$$\begin{aligned} \mathcal{S}[P,Q] &= -2 \int P(x'|x) \log \frac{P(x'|x)}{\gamma^{1/2}(x')} + 2 \int dx' P(x'|x) S(x') \\ &\quad - \frac{1}{2} \int P(x'|x) \gamma_{ab}(x) \Delta x^a \Delta x^2 - \int dx' P(x'|x) A'_a(x) \Delta x^a \,, \end{aligned} \tag{105}$$

where[1]

$$A'_a(x) \stackrel{\text{def}}{=} A_a(x) - F_a(x) \,. \tag{106}$$

Up to an irrelevant proportionality constant, the joint entropy (105) can be written as

$$\begin{aligned} \mathcal{S}[P,Q] &= - \int P(x'|x) \log \frac{P(x'|x)}{\gamma^{1/2}(x')} + \int dx' P(x'|x) S(x') \\ &\quad - \frac{1}{4} \Delta \bar{\ell}^2 - \frac{1}{2} \int dx' P(x'|x) A_a(x) \Delta x^a \,, \end{aligned} \tag{107}$$

where

$$\Delta \bar{\ell}^2 = \int P(x'|x) \gamma_{ab}(x) \Delta x^a \Delta x^2 \,. \tag{108}$$

[1]Here however A'_a, A_a and F_a are abstract objects. However if A_a is identified with the vector potential, then eq. (106) involves some sort of gauge transformation. Recall the gauge transformation $A_a \to A'_a = A_a + \partial_a \chi$ so that $A'_a \to A_a = A'_a - \partial_a \chi$. In our case A_a, F_a and A'_a are given by eqs. (95), (103) and (106) respectively. Therefore $\chi = \int dx'' P(x''|x) S(x'')$.

Next we vary $P(x'|x)$ to maximize $\mathcal{S}[P,Q]$ subject to the constraints (74), (75), and (76) but there is one last constraint missing.
The forth constraint – The normalization condition:

$$\int dx' P(x'|x) = 1\,. \tag{109}$$

Note that for notational convenience the measure d^3x' is written as dx'.

Before employing the ME method, for later convenience write eq. (107) as

$$\begin{aligned}\mathcal{S}[P,Q] &= -\int P(x'|x)\log\frac{P(x'|x)}{\gamma^{1/2}(x')} + \int dx' P(x'|x) S(x') \\ &\quad -\frac{\alpha}{2}\Delta\bar{\ell}^2 + \lambda \int dx' P(x'|x) A_a(x)\Delta x^a\,,\end{aligned} \tag{110}$$

where

$$\alpha = \frac{1}{2} \quad \text{and} \quad \lambda = -\frac{1}{2}\,. \tag{111}$$

Now use the machinery of ME,

$$\delta[\mathcal{S}[P,Q] - \alpha_0 \textstyle\int dx' P(x'|x)] = 0\,. \tag{112}$$

The result is

$$P(x'|x) = \frac{1}{\zeta(x,\alpha,\lambda)}\gamma^{1/2}(x')\exp\left[S(x') - \tfrac{\alpha}{2}\gamma_{ab}(x)\Delta x^a\Delta x^b - \lambda A_a(x)\Delta x^2\right]\,, \tag{113}$$

where

$$\zeta(x,\alpha,\lambda) = \int dx' \gamma^{1/2}(x')\exp\left[S(x') - \tfrac{\alpha}{2}\gamma_{ab}(x)\Delta x^a\Delta x^b - \lambda A_a(x)\Delta x^2\right]\,. \tag{114}$$

The transition probability (113) is meant to hold for short steps but this is not obvious. Technically there are two difficulties:

1. First the constant α is not large rather it is given by (111). Therefore eq. (113) does not lead to the Brownian motion. In Ref. [4] α is a Lagrange multiplier so it is possible to derive diffusion process from (21). Here we do not have this freedom. The constants α and λ are not Lagrange multipliers. They are introduced just for convenience.

2. Second the space X is not assumed to be flat. In a curved space, the displacement Δx^a does not transform like a vector. The second order effects must be taken into account. However this difficulty can be resolved. In the author's PhD thesis [15] entropic dynamics is successfully extended to curved spaces.

But we are interested to analyze eq. (113) in any way. Let us for the time being insist that α is not given by eq. (111), and assume that α is very large. Having assumed that α is large, let us write eq. (113) in the following form

$$P(x'|x) = \gamma^{1/2}(x')P'(x'|x)\,, \tag{115}$$

with

$$P'(x'|x) = \frac{1}{\zeta(x,\alpha,\lambda)} \exp\left[S(x') - \tfrac{\alpha}{2}\gamma_{ab}(x)\Delta x^a \Delta x^b - \lambda A_a(x)\Delta x^2\right], \tag{116}$$

In eq. (115), $P(x'|x)$ is tensor of rank zero, its transformation involves a Jacobian factor $\gamma^{1/2}(x')$. On the other hand $P'(x'|x)$ is an invariant scalar density, it transformation does not involve any Jacobian.

All we need to examine $\mathcal{P}(x'|x)$. We have already required that α is large, therefore $P(x'|x)$ holds for short steps. Now $P'(x'|x)$ could be easily expressed as a Gaussian if the metric $\gamma_{ab}(x)$ were that of the flat space, or otherwise write $P'(x'|x)$ in a locally Cartesian coordinate, also called the *normal coordinate*. In normal coordinates (NC) at a point p the metric tensor in the vicinity of p is approximately that of flat Euclidean space. That is, if p is a point in the manifold then

$$\gamma_{ab}(x_p) = \frac{\delta_{ab}}{\sigma^2}\,, \tag{117}$$

so that

$$\left.\frac{\partial \gamma_{ab}}{\partial x^c}\right|_{x_p} = 0\,, \tag{118}$$

however

$$\left.\frac{\partial^2 \gamma_{ab}}{\partial x^c \partial x^d}\right|_{x_p} \neq 0\,, \tag{119}$$

they are the effects of curvature if the manifold is not exactly flat.

For large α, expanding the exponent of eq. (116)) about its maximum.

$$P'(x'|x) \approx \frac{1}{Z(x)} \exp\left[-\frac{\alpha(x)}{2\sigma^2}\delta_{ab}\left(\Delta x^a - \Delta \bar{x}^a\right)\left(\Delta x^b - \Delta \bar{x}^b\right)\right]. \tag{120}$$

This is the expression for transition probability in normal coordinates which is obviously a Gaussian. The factors independent of x' are absorbed into a new normalization $Z(x)$. The displacement Δx^a and the expected drift $\Delta \bar{x}^a$ are given by

$$\Delta x^a = \Delta \bar{x}^a + \Delta w^a\,, \tag{121}$$

$$\Delta \bar{x}^a = \frac{\sigma^2}{\alpha}\delta^{ab}[\partial_b S(x) - \lambda\, \mathcal{A}_a(x)]\,, \tag{122}$$

and the fluctuations

$$\langle \Delta w^a \rangle = 0\,, \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{\sigma^2}{\alpha}\delta^{ab}\,. \tag{123}$$

For large α the step size becomes very small but the fluctuations become dominant, because $\Delta \bar{x} \sim O(\alpha^{-1})$ while $\Delta w \sim O(\alpha^{-1/2})$. It means that as $\alpha \to \infty$ the trajectory is continuous but not differentiable—just like Brownian motion.

Now restoring (111) so that $\alpha = 1/2$, then up to a proportionality constant σ^2,

$$\Delta x^a > 1\,. \tag{124}$$

But this ruins everything. Recall that the transition probability (113) followed by making the Taylor expansion of $P(x''|x', x)$ about x while assuming eq. (84),

$$\Delta P = \frac{\partial P}{\partial x^{a_1}} \Delta x^{a_1} + \frac{1}{2} \frac{\partial^2 P}{\partial x^{a_1} \partial x^{a_2}} \Delta x^{a_1} \Delta x^{a_2} + \cdots , \tag{125}$$

where on the r. h. s. $\Delta P = P(x''|x', x) - P(x''|x)$ and on the r. h. s. $P = P(x''|x)$. Therefore ΔP grows up when $\Delta x^{a_1} > 1$ and hence eq. (113) does not follow.

6 The Renormalization

We observed that Δx^a does not behave like a differential. To remedy this serious problem we change the role of *differential* and *derivative*.

Define

$$\mathcal{D}x_{a_1} \stackrel{\text{def}}{=} \frac{\partial P}{\partial x^{a_1}} . \tag{126}$$

Equation (126) is very unusual. On the l. h. s. is some differential quantity while the r. h. s. is the derivative. So the differential is equal to the derivative! To make sense of it we further need to clean up our notations.

In fact eq. (126) is some sort of a *duality* equation. On the l. h. s. is an object that is an element of something we call it *category*[2], while r. h. s. involves an object of another category. We also need to name these categories. Let us start with the r. h. s. The object on the r. h. s. is an element of **G** or **math**. It is the category of all objects in the standard mathematics. The object on the l. h. s. is an element of $\tilde{\mathbf{G}}$, or **comath**. It is the category of co-mathematics.

Since (126) involves objects of different mathematics', we shall write it as

$$\mathcal{D}x_{a_1}|_{\tilde{\mathbf{G}}} = \left.\frac{\partial P}{\partial x^{a_1}}\right|_{\mathbf{G}} , \tag{127}$$

or simply drop the vertical bars and reserve Calagraphic symbols for the objects in **comath**. However if the Calagraphic symbols are not convenient then use the vertical bars.

Now we inspect $\Delta x = x' - x$. We know that x is the earlier point and x' is the later point. It means that Δx is something that takes into account two different points in space. On the other hand conditional probability is also something that has this property. From this we conclude that

$$\left.\int \mathcal{P}^{a_1}(x'|x)\right|_{\tilde{\mathbf{G}}} = \Delta x^{a_1}|_{\mathbf{G}} \tag{128}$$

where $\mathcal{P}^a$ is the conditional probability in **comath**. It is a tensor of rank 1. The l. h. s. of eq. (128) represents an integral operator but it work like a derivative in **comath**, see section 7. Therefore

$$\left.\frac{\partial}{\partial x}\right|_{\tilde{\mathbf{G}}} = \left.\int dx\right|_{\mathbf{G}} \tag{129}$$

[2] The category in standard mathematics might have different meanings.

Combining eqs. (127) and (128), we have

$$\int \mathcal{P}^{a_1}(x'|x)\,\mathcal{D}x_{a_1}\Big|_{\tilde{\mathbf{G}}} = \frac{\partial P}{\partial x^{a_1}}\Delta x^{a_1}\Big|_{\mathbf{G}} . \tag{130}$$

Similarly

$$\int\int \mathcal{P}^{a_1 a_2}(x'|x)\mathcal{D}x_{a_1}\mathcal{D}x_{a_2}\Big|_{\tilde{\mathbf{G}}} = \frac{\partial^2 P}{\partial x^{a_1}\partial x_{a_2}}\Delta x^{a_1}\Delta x^{a_1}\Big|_{\mathbf{G}} , \tag{131}$$

where $\mathcal{P}^{a_1 a_2}$ is a rank 2 conditional probability in **comath**. With the same procedure the higher order terms series 125 can be constructed.

7 The Duality Principle

Without proof let me state the followings. With each of the following statements *vise versa* is understood. We shall call these statements as the *duality principle* (DP).

1. The rules of mathematics that hold in **math** they continue to hold in **comath**. However their roles might reverse.

2. The conditional probabilities in **math** becomes functions in **comath**.

$$\mathcal{F}(x)|_{\tilde{\mathbf{G}}} = P(x'|x)|_{\mathbf{G}} . \tag{132}$$

   Note that $\mathcal{F}$ is a function of the earlier point x.

3. The zeros in **math** become infinities in **comath**.

And so on.

And yet the categories **math** and **comath** are subcategories of a larger category **H**. The category **H** will be discussed elsewhere.

8 The Duality Operator

Let $\hat{\mathbf{D}}$ be an operator. We shall call it the duality operator. It has the following properties:

1. Let $\mathcal{A}$ be an object in **comath** and A be its dual in **math**, then

$$\hat{\mathbf{D}}\mathcal{A} = A , \tag{133}$$

   so that[3]

$$\hat{\mathbf{D}}\hat{\mathbf{D}}\mathcal{A} = \hat{\mathbf{D}}A = \mathcal{A} . \tag{134}$$

[3]This operator might working differently if several subcategories of **H** are considered. Here we are only dealing with two dual subcategories of **H**.

2.
$$\hat{\mathbf{D}}[\ =\ ]_{\tilde{\mathbf{G}}} = [\ =\ ]_{\mathbf{G}}\,. \tag{135}$$

The equality sign '=' is its dual.

The inequalities,

$$\hat{\mathbf{D}}[\ <\ ]_{\tilde{\mathbf{G}}} = [\ >\ ]_{\mathbf{G}} \ \text{ and vice versa.} \tag{136}$$

3. **Constant function:** Beside conditional probabilities there are also unconditional probabilities such as $P(x)$, then

$$\hat{\mathbf{D}}\,[\ P(x)\ ]_{\mathbf{G}} = [\ \mathcal{C}\ ]_{\tilde{\mathbf{G}}}\ , \tag{137}$$

where $\mathcal{C}$ is some constant function in **comath**. The *constants* in **math** are *probabilities* in **comath** and vice versa.

4. **Conditional probability of several arguments:** Let $P(x_1, x_2|x_3, x_4)$ be a conditional probability in **math**, where x_1, x_2, x_3 and x_4 are all distinct points. Then by 132,

$$\hat{\mathbf{D}}\,[P(x_1, x_2|x_3, x_4)]_{\mathbf{G}} = [\ \mathcal{F}(x_3, x_4)\ ]_{\tilde{\mathbf{G}}}\ . \tag{138}$$

But this does not make sense. A function[4] is defined at a unique point in space. It does not take into account different points of space. It is the conditional probability (distributions) that take into account several points in space. To make sense of it then it must be the case that x_3 and x_4 are one and the same point. But they are given to be distinct points. Before we reach a conclusion we generalize eq. (138).

Let y be in Y, and let I be an interval in the real numbers R. The interval I could be the entire R, then

$$\hat{\mathbf{D}}\,[P(y|I)]_{\mathbf{G}} = [\ \mathcal{F}(I)\ ]_{\tilde{\mathbf{G}}}\ . \tag{139}$$

For $\mathcal{F}$ to be a legitimate function, then it must the case that

$$[I = x]_{\tilde{\mathbf{G}}}\ . \tag{140}$$

This means that the universe in **math** becomes a *dot* in **comath**. If x is a point then what is its value? In fact x is arbitrary it can have any value as you wish. Different values of it corresponds to different *solutions*. We finally write eq. (140) in the following convenient form

$$[x = \ \bullet\ ]_{\tilde{\mathbf{G}}} \tag{141}$$

[4]Just for the convenience of some readers a *function* is something that is defined on a point say x. Let $f(x)$ be a function, then $f(x_0)$ is the value of f at the point x_0.

9 More on the math and the comath

We wish to further clarify our notations. When we write say

$$[A = B] \ . \tag{142}$$

Note the outer square brackets. The outer brackets encloses say an equation. The whole object including the outer brackets and the equation inside is a *box*. The box is not an equation. It has no left and right hand sides. It is just a box.

When the duality operator is applied to the box, then the operation should be performed very carefully,

$$\hat{\mathbf{D}}\,[A = B] \ , \tag{143}$$

The operator $\hat{\mathbf{D}}$ has to pass through each object in the box *linearly*. And the final result should not be written in the same line otherwise it may create confusions. A better way of writings might be the followings:

$$\begin{gathered} \hat{\mathbf{D}}[A = B] \\ \Rightarrow \\ [\ [\hat{\mathbf{D}}A]\ [\hat{\mathbf{D}} =]\ [\hat{\mathbf{D}}B]\ ] \\ \Rightarrow \\ [\mathcal{A} = \mathcal{B}] \end{gathered} \tag{144}$$

And thus we ended up with a box.

Additive Inverse: Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be in **comath**, then

$$\begin{gathered} [\mathcal{A}_1 - \mathcal{A}_1 = \bullet_{c_{\mathcal{A}_1}}]_{\tilde{\mathbf{G}}} \\ \Rightarrow \\ [\mathcal{A}_1 + \bullet_{c_{\mathcal{A}_1}} = \mathcal{A}_1]_{\tilde{\mathbf{G}}} \end{gathered} \tag{145}$$

and

$$\begin{gathered} [\mathcal{A}_2 - \mathcal{A}_2 = \bullet_{c_{\mathcal{A}_2}}]_{\tilde{\mathbf{G}}} \\ \Rightarrow \\ [\mathcal{A}_2 + \bullet_{c_{\mathcal{A}_2}} = \mathcal{A}_2]_{\tilde{\mathbf{G}}} \end{gathered} \tag{146}$$

where $c_{\mathcal{A}_1}$ and $c_{\mathcal{A}_2}$ are the color indices. Different dots have different colors.

The elements $\mathcal{A}_1$ and $\mathcal{A}_2$ etc. are in fact all dots in **comath**. Also note that within **comath** the operations of say $+$, $-$, $\times$ and $\div$ are similar to that of their operations in **math**. However they flip to their duals under the duality transformation. For example,

$$[\ [\hat{\mathbf{D}}+]_{\tilde{\mathbf{G}}} = [-]_{\mathbf{G}}\ ] \quad \text{etc.} \tag{147}$$

But within **comath**,

$$\begin{gathered} [\ [\mathcal{A} + \mathcal{B} = \mathcal{C}]_{\tilde{\mathbf{G}}}\ ] \\ \Rightarrow \\ [\ [\ \mathcal{A} + \mathcal{B} - \mathcal{C} = \bullet_{\mathcal{C}}\ ]_{\tilde{\mathbf{G}}}\ ] \quad \text{etc.} \end{gathered} \tag{148}$$

We now return to the Taylor series (125)

$$\left[+P(x''|x',x) - P(x'|x) = +\frac{1}{1!}\frac{\partial P(x''|x)}{\partial x^{a_1}}\Delta x^{a_1} + \frac{1}{2!}\frac{\partial^2 P(x''|x)}{\partial x^{a_1}\partial x^{a_2}}\Delta x^{a_1}\Delta x^{a_2} + \cdots\right]_{\mathbf{G}}, \tag{149}$$

Apply the operator $\hat{\mathbf{D}}$ to box (149)

$$\hat{\mathbf{D}}\left[+P(x''|x',x) - P(x'|x) = +\tfrac{1}{1!}\tfrac{\partial P(x''|x)}{\partial x^{a_1}}\Delta x^{a_1} + \tfrac{1}{2!}\tfrac{\partial^2 P(x''|x)}{\partial x^{a_1}\partial x^{a_2}}\Delta x^{a_1}\Delta x^{a_2} + \cdots\right]_{\mathbf{G}}$$
$$\Rightarrow$$
$$\left[-\mathcal{F}(x) + \mathcal{F}(x) = - \int \mathcal{P}_{(1/1!)}(x) P^{a_1}(x''|x)\mathcal{D}x_{a_1} - \int\int \mathcal{P}_{(1/2!)}(x) P^{a_1 a_2}(x''|x)\mathcal{D}x_{a_1}\mathcal{D}x_{a_1} - \cdots\right]_{\tilde{\mathbf{G}}}$$
$$\Rightarrow$$
$$[\ [-][\ ] = [-][\ ]\ ]_{\tilde{\mathbf{G}}}$$
$$\Rightarrow$$
$$\left[\bullet_{\mathcal{F}(x)} = \int \mathcal{P}_{(1/1!)}(x) P^{a_1}(x''|x)\mathcal{D}x_{a_1} + \int\int \mathcal{P}_{(1/2!)}(x) P^{a_1 a_2}(x''|x)\mathcal{D}x_{a_1}\mathcal{D}x_{a_2} + \cdots\right]_{\tilde{\mathbf{G}}} \tag{150}$$

This is the Taylor series of $\mathcal{F}(x)$ in **comath** . Let us call it the co-Taylor series of $\mathcal{F}(x)$. Like box (144), the operator $\hat{\mathbf{D}}$ had to pass linearly through each of the object inside the box. Here we have used various results obtained previously. For example, by (137)

$$[\ \hat{\mathbf{D}}[1/2!]_{\mathbf{G}} = [\mathcal{P}_{1/2!}(x)]_{\tilde{\mathbf{G}}}\ ] \tag{151}$$

In fact $P^{a_1}(x''|x)$ is a delta function in **comath** of the type

$$[P^{a_1}(x''|x) = \Delta^{a_1}(x''-x)]_{\tilde{\mathbf{G}}}\,, \tag{152}$$

and similarly

$$[P^{a_1 a_2}(x''|x) = \Delta^{a_1 a_2}(x''-x)]_{\tilde{\mathbf{G}}}\,. \tag{153}$$

It has the characteristics of both the Kronecker delta and the Dirac delta because in **math** the Kronecker delta carries the indices say $a_1 a_2$ while the Dirac delta has the argument $x'' - x$. The co-Delta in **comath** unifies the Kronecker and Dirac delta's.

But the single index of the co-Delta in eq. (152) makes it complicated to deal with the indices. So we write it in the following convenient form

$$[\Delta^{a_1}(x''-x) = \epsilon^{a_1 a_2 a_3}\delta_{a_2 a_3}\delta(x''-x)]_{\tilde{\mathbf{G}}}\,, \tag{154}$$

and

$$[\Delta^{a_1 a_2}(x''-x) = \epsilon^{a_1 a_2 a_3 a_4}\delta_{a_3 a_4}\delta(x''-x)]_{\tilde{\mathbf{G}}}\,, \tag{155}$$

it involves all three the Levi-Civita symbol, the Kronecker delta and the Dirac delta. The Levi-Civita symbol is symmetric, the Kronecker delta is anti-symmetric while Dirac delta also reverses its role,

$$\left[\delta(x''-x) = \begin{cases} \bullet_0 & \text{if } x'' = x \\ \bullet_\infty & \text{if } x'' \neq x \end{cases}\right]_{\tilde{\mathbf{G}}}, \tag{156}$$

then

$$\left[\int \delta(x'' - x) = \bullet_\infty\right]_{\tilde{\mathbf{G}}} . \tag{157}$$

From it follows,

$$\left[\int \Delta^{a_1}(x'' - x) = \bullet_0\right]_{\tilde{\mathbf{G}}} , \tag{158}$$

and

$$\left[\int \mathcal{P}(x)\Delta^{a_1}(x'' - x) = \mathcal{P}^{a_1}(x'')\right]_{\tilde{\mathbf{G}}} . \tag{159}$$

10 The Physics

Here we apply it to the problems in physics. As an example here we obtain an exact solution of SE for an arbitrary potential. We note that the dual categories **math** and **comath** are abstract. One can attach any meanings to it depending on the problem. For physics we identify **math** with the physical universe while **comath** is identified with the *dark universe* (dark matter and dark energy).

First we analyze the free particle case. The representation of SE in the physical universe is given by[5]

$$\left[c_1 \frac{\partial \Psi(z)}{\partial t} = c_2 \nabla^2 \psi(z)\right]_{\mathrm{G}} , \tag{160}$$

where $c_1 = i\hbar$, and $c_2 = -\hbar^2/2m$, and $z = (x, t)$. Therefore the representation of SE in the dark universe becomes.
(see next page)

[5]Here we consider a single particle, however, we arrive at the same conclusions if several particles are considered.

$$
\begin{array}{c}
\left[\int \mathcal{P}_{c_1}(t)\mathcal{P}^t(t'|t) = \int \mathcal{P}_{c_2}(x)\mathcal{P}_i^i(x'|x)\right]_{\tilde{\mathbf{G}}} \\
\Rightarrow \\
\left[\int \mathcal{P}_{c_1}(t)\Delta^t(t'-t) = \int \mathcal{P}_{c_2}(x)\Delta_i^i(x'-x)\right]_{\tilde{\mathbf{G}}} \\
\Rightarrow \\
\left[\mathcal{P}_{c_1}^t(t') = \mathcal{P}_{c_2\ \ i}^{i}(x')\right]_{\tilde{\mathbf{G}}} \\
\Rightarrow \\
\hat{\mathbf{D}}\left[\mathcal{P}_{c_1}^t(t') = \mathcal{P}_{c_2\ \ i}^{i}(x')\right]_{\tilde{\mathbf{G}}} \\
\Rightarrow \\
\left[c_1^t = c_{2\ \ i}^{i}\right]_{\mathrm{G}} \\
\Rightarrow \\
\left[i^t\hbar = -\frac{\hbar^2}{2m}\right]_{\mathrm{G}} \\
\Rightarrow \\
\left[m = \frac{\hbar}{2}i^t\right]_{\mathrm{G}}
\end{array}
\qquad (161)
$$

The tensorial nature of box (161) is not preserved because it is the non relativistic case. But this is a strange result. The mass is purely imaginary! Here i^t is the time component of i. Since it shows that the mass is constant, we suspect that m might be the mass of the entire universe.

To further explore it we need to obtain the exact solution SE for an arbitrary potential,

$$
\begin{array}{c}
\left[c_1\frac{\partial\Psi(z)}{\partial t} = c_2\nabla^2\psi(z) + V(z)\Psi(z)\right]_{\mathrm{G}} \\
\Rightarrow \\
\left[\int \mathcal{P}_{c_1}(t)\mathcal{P}_{\Psi}^t(t'|t) = \int \mathcal{P}_{c_2}(x)\mathcal{P}_{\Psi\ \ i}^{j}(x'|x) + \mathcal{P}_V(z'|z)\mathcal{P}_{\Psi}(z'|z)\right]_{\tilde{\mathrm{G}}} \\
\Rightarrow \\
\left[\int \mathcal{P}_{c_1}(t)\Delta^t(t'-t) = \int \mathcal{P}_{c_2}(x)\Delta_i^i(x'-x) + \Delta(z'-z)\mathcal{P}_{\Psi}(z'|z)\right]_{\tilde{\mathrm{G}}} \\
\Rightarrow \\
\left[\mathcal{P}_{c_1}^t(t') = \mathcal{P}_{c_2\ \ i}^{i}(x') + \bullet_{\Psi}\right]_{\tilde{\mathrm{G}}} \\
\Rightarrow \\
\left[\int \mathcal{P}_{c_1}(t)\Delta^t(t'-t) = \int \mathcal{P}_{c_2}(x)\Delta_i^i(x'-x) + \Delta(z'-z)\mathcal{P}_{\Psi}(z|z)\right]_{\tilde{\mathrm{G}}} \\
\Rightarrow \\
\left[\mathcal{P}_{c_1}^t(t') = \mathcal{P}_{c_2\ \ i}^{i}(x') + \bullet_{\Psi}\right]_{\tilde{\mathrm{G}}}
\end{array}
\qquad (162)
$$

Since $\mathcal{P}_{c_1}^t(t')$ and $\mathcal{P}_{c_2\ \ i}^{i}(x')$ are probabilities, therefore the dot is probabilistic in the dark universe.

We now want to transform the solution to the physical universe by applying the duality operator.

$$\hat{\mathrm{D}}\left[\mathcal{P}^t_{c_1}(t') = \mathcal{P}^i_{c_2\ i}(x') + \bullet_\Psi\right]_{\tilde{\mathrm{G}}}$$
$$\Rightarrow$$
$$\left[c^t_1 = c^i_{2\ i} + \Psi(\bullet)\right]_{\tilde{\mathrm{G}}} \tag{163}$$

Previously we have $c^t_1 = c^i_{2\ i}$. Therefore

$$[\Psi(\bullet) = 0]_{\mathrm{G}} \ , \tag{164}$$

From box (141) the value of dot is arbitrary, therefore

$$[\Psi(x) = 0, \quad \text{for all } x]_{\mathrm{G}} \ , \tag{165}$$

Therefore the wave function is zero in one universe while it is probabilistic in the dual universe.
summary:
There are no particles. The entire universe is a compact dot that is probabilistic in one universe (dark matter and dark energy) while its wave function is zero in the dual universe (physical universe). Furthermore the fundamental constants of nature are constant in one universe while they are probabilities in the dual universe.

11 Conclusions

In this paper we have considered only two subcategories of **H**. We might obtain more interesting results if more elements of **H** or rather the entire **H** is considered. Secondly the conclusions of the paper shows that probability plays an important role. An in another paper [16] we have qualitatively shown that probability theory is a weaker form of the *question theory* (QT). We expect that the TOE's of QT might be more interesting.

References

[1] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Physical Review* **106** (1957) 620.

[2] E. T. Jaynes, "Information Theory and Statistical Mechanics. II," *Physical Review* **108** (1957) 171.

[3] A. Caticha, *Entropic Inference and the Foundations of Physics.* (São Paulo, Brazil, 2012). `http://www.albany.edu/physics/ACaticha-EIFP-book.pdf`.

[4] A. Caticha, "Entropic Dynamics, Time and Quantum Theory," *Journal of Physics A Mathematical General* **44** (2011) 225303, `arXiv:1005.2357 [quant-ph]`.

[5] J. W. Gibbs, *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics.* (Dover Publication, New York, 1960).

[6] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* (The University of Illinois Press, Urbana, 1964).

[7] A. Caticha, "Relative Entropy and Inductive Inference," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, G. J. Erickson and Y. Zhai, eds., vol. 707 of *American Institute of Physics Conference Series*, p. 75. 2004. arXiv:physics/0311093.

[8] A. Caticha, "Information and Entropy," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, K. H. Knuth, A. Caticha, A. Giffin, and C. C. Rodríguez, eds., vol. 954 of *American Institute of Physics Conference Series*, p. 11. 2007. arXiv:0710.1068 [physics.data-an].

[9] A. Caticha, "Change, Time and Information Geometry," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, A. Mohammad-Djafari, ed., vol. 568 of *American Institute of Physics Conference Series*, p. 72. 2001. arXiv:math-ph/0008018.

[10] A. Caticha, "Entropic Time," in *American Institute of Physics Conference Series*, A. Mohammad-Djafari, J.-F. Bercher, and P. Bessiére, eds., vol. 1305 of *American Institute of Physics Conference Series*, p. 200. 2011. arXiv:1011.0746 [quant-ph].

[11] S. Chandrasekhar, "Stochastic Problems in Physics and Astronomy," *Rev. Mod. Phys.* **15** (1943) 1. http://link.aps.org/doi/10.1103/RevModPhys.15.1.

[12] E. Nelson, *Quantum Fluctuations.* (Princeton U. Press, Princeton, 1985).

[13] E. Nelson, "Connection between Brownian Motion and Quantum Mechanics," in *Einstein Symposion Berlin*, H. Nelkowski, A. Hermann, H. Poser, R. Schrader, and R. Seiler, eds., vol. 100 of *Lecture Notes in Physics, Berlin Springer Verlag*, p. 168. 1979.

[14] S. Nawaz and A. Caticha, "Momentum and Uncertainty Relations in the Entropic Approach to Quantum Theory," in *American Institute of Physics Conference Series*, P. Goyal, A. Giffin, K. H. Knuth, and E. Vrscay, eds., vol. 1443 of *American Institute of Physics Conference Series*, p. 112. 2012. arXiv:1108.2629 [quant-ph].

[15] S. Nawaz, *Momentum and Spin in Entropic Quantum Dynamics.* PhD thesis, State University of New York at Albany, 2013.

[16] S. Nawaz, "The Topological Structure of Question Theory," *ArXiv e-prints* (2010) , arXiv:1006.2481 [math.GM].