Maximizing the Entropy under Geometric Constraints: A Simple and Plausible Universal Foundation for Physics

Alexandre Harvey-Tremblay[†]

[†]Independent scientist, aht@protonmail.ch
July 17, 2022

Abstract

In this study, we introduced the notion of a geometric constraint and we used it to derive the probability measure that maximizes the entropy subject to this constraint. We obtained a measure that admits an algebra of observables spawning all possible geometric measurements in nature. Gravity followed very naturally from the measure in the form of a gauge theory of (up to) the affine group. We further show that the measure, when demanding that it preserves the Dirac current, accepts only two categories of geometric observables: one evolving according to the SU(2) $\times$ U(1) gauge symmetry and the other according to the SU(3) gauge symmetry. To construct the constraint, the key idea is to connect geometry with the theory of probability by using the trace. The trace of a matrix can be seen as the expected eigenvalue times the dimension of the vector space, and the eigenvalues are the ratios of the distortion of the geometric transformation associated with the matrix. Results regarding the foundations of quantum mechanics are also obtained: a plausible origin for the Born rule (including the wave-function) is revealed from the union of geometry and entropy; also, we find the wave-function collapse problem to be superseded by a theory of geometric measurements, satisfying the axioms of quantum mechanics and inherited from statistical mechanics, which we introduce as the metrological interpretation.

Keywords: Gravity, quantum physics, standard model, geometric constraint

1 Introduction

A new form of constraint referred to as the *geometric constraint* is introduced. This constraint extends the tools of statistical mechanics to geometric and quantum systems.

Using this constraint, the entropy is maximized to produce a geometric probability measure. The measure supports gravity as a affine gauge theory in any dimensions, and in four-dimensions (4D) accepts two observables to preserve the Dirac current: one having the SU(2) × U(1) gauge symmetry and the other having the SU(3) gauge symmetry. This makes the method particularly interesting because it accepts a notion of particle physics in addition to gravity. Let us state that have not yet attempted to quantized the resulting gravitational theory; here we hope to lay the foundation of the entropy maximization process, the geometric constraints and to showcase the advantages of this framework.

The key idea is to connect geometry and the probability theory using the trace. The trace accepts a probability interpretation[1] as the expectation value of the eigenvalues times the dimension of the vector space. It also connects to the geometry as the eigenvalues are the ratio of the distortion of the geometric transformation associated with the matrix.

The geometric constraint is defined as

$$\operatorname{tr}\overline{\mathbf{M}} = \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr}\mathbf{M}(q), \tag{1}$$

where $\mathbf{M}$ is an arbitrary $n \times n$ matrix, and $\mathbb{Q}$ is a statistical ensemble. Here, $\operatorname{tr}\overline{\mathbf{M}}$ denotes the expectation value of the statistically weighted sum of matrices $\mathbf{M}(q)$ parametrized over the ensemble $\mathbb{Q}$.

Alternatively (and preferably), we may use the geometric algebra to define the constraint. We will use this approach in this paper. In this case, it will be defined as

$$\operatorname{tr}\overline{\mathbf{u}} = \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr}\mathbf{u}(q), \tag{2}$$

where $\mathbf{u}$ is an arbitrary multi-vector of the real geometric algebra in n dimensions $\mathbb{G}(n, \mathbb{R})$. In either case, the constraints are equally expressive, but the use of multi-vectors rather than matrices makes the geometric character of the method stand out. More details on geometric algebra (and the present notation) are provided in the method section.

In statistical mechanics, using this equality as a constraint on the entropy is a claim that we can observe (up to a phase) the distortions produced by any geometric transformations in nature and that the permissible statistics preserve the expectation value of these distortions. For instance, a statistical system measured exclusively using a ruler, clock, and protractor will carry, following our entropy maximization procedure, the Lorentz group symmetry in its associated probability measure.

In statistical mechanics, constraints are used to derive the Gibbs measure using Lagrange multipliers[2] by maximizing the entropy.

For instance, an energy constraint on the entropy is

$$\overline{E} = \sum_{q\in\mathbb{Q}} \rho(q) E(q), \tag{3}$$

which is associated with an energy meter measuring the system energy and producing a series of energy measurements $E_1, E_2, \ldots$ converging to an expectation value $\overline{E}$.

Another common constraint is that of the volume

$$\overline{V} = \sum_{q \in \mathbb{Q}} \rho(q) V(q), \tag{4}$$

which is associated with a volume meter acting on the system by producing a sequence of measurements of the volume $V_1, V_2, \ldots$ converging to an expectation value $\overline{V}$.

Moreover, the sum over the statistical ensemble must be equal to 1, as shown below.

$$1 = \sum_{q \in \mathbb{Q}} \rho(q) \tag{5}$$

With equations (3) and (5), the typical system of statistical mechanics is obtained by maximizing the entropy using its corresponding Lagrange equation. The Lagrange multipliers method is expressed as

$$\mathcal{L} = -k_B \sum_{q \in \mathbb{Q}} \rho(q) \ln \rho(q) + \lambda \left(1 - \sum_{q \in \mathbb{Q}} \rho(q)\right) + \beta \left(\overline{E} - \sum_{q \in \mathbb{Q}} \rho(q) E(q)\right), \tag{6}$$

where, λ and β are Lagrange multipliers.

Therefore, solving $\frac{\partial \mathcal{L}}{\partial \rho} = 0$ for ρ, we obtain the Gibbs measure as

$$\rho(q, \beta) = \frac{1}{Z(\beta)} \exp(-\beta E(q)), \tag{7}$$

where,

$$Z(\beta) = \sum_{q \in \mathbb{Q}} \exp(-\beta E(q)). \tag{8}$$

In our procedure, we replace (3) with $\operatorname{tr} \overline{\mathbf{M}}$, and the constraint is now geometric. Instead of energy meters or volume meters, we have rulers, clocks, protractors, spin meters, dilation meters, and shear meters. This set of instruments may exceed what is constructible in nature (e.g. dilation or shear meters), however facing the unknown we prefer to develop the theory in its full generality.

For our procedure to properly connect to quantum mechanics, the statistical interpretation of the entropy must be altered with respect to its statistical mechanics interpretation. The probability measure will be interpreted as quantifying the information associated with the receipt of a message of measurements. Therefore, we replace the Boltzmann entropy with the Shannon entropy. This replacement does not change the form of the mathematical equation for the entropy (the expressions for the Boltzmann and the Shannon entropies are the same up to a multiplication constant) but only the final interpretation (this will be further detailed in the discussion, section 6).

The corresponding Lagrange equation is

$$\mathcal{L} = -\sum_{q\in\mathbb{Q}} \rho(q) \ln \rho(q) + \lambda \left(1 - \sum_{q\in\mathbb{Q}} \rho(q)\right) + \tau \left(\operatorname{tr} \overline{\mathbf{u}} - \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{u}(q)\right), \quad (9)$$

and it is now sufficient to solve $\frac{\partial \mathcal{L}}{\partial \rho} = 0$ for ρ to obtain the solution.

The manuscript is organized as follows: In the methods section, referencing the work of Lundholm[3], we will introduce a number of tools using geometric algebra. Specifically, we will introduce the notion of a determinant for multivectors and notions of a Clifford conjugate generalizing the complex conjugate. These tools will allow us to entirely express our results geometrically.

In the results section, we will present two solutions to the Lagrange equation. The first is a recovery of standard non-relativistic quantum mechanics, which occurs when the matrix is reduced from an arbitrary matrix to a representation of the imaginary number. The second is the general case with an arbitrary matrix or multi-vector.

We then develop our initial results into a geometric foundation to physics, both in two-dimensional (2D) and 4D consistent with the general solution. In 2D, the self-adjoint observables are generalized to observables equal to their Clifford conjugate. Remarkably, in 4D, we obtain an even more sophisticated relation for observables pitting four terms, which together satisfy the SU(2) $\times$ U(1) and the SU(3) gauge symmetry. Lastly for this section, we discuss the prospects of a gauge theory of gravity, which exploits the flexibility of our probability measure to remain normalizable and invariant with respect to all general linear transformation (and superposition thereof) which be believe are required to accommodate gravity in 4D.

Finally, in the discussion, we introduce an interpretation of quantum mechanics consistent with its newly revealed origin as the measure maximizing the Shannon entropy subject to constrainment by geometric measurements, which we call the metrological interpretation. In this interpretation, the measurements and the constraint they entail on the entropy are considered more fundamental than the wave function which is entirely derivable from them. The end product is a theory which deprecates the measurement problem, superseding it with theory of instrumentation, and provides a plausible account for the origins of quantum mechanics in nature tying it to the geometric measurements that are possible.

2 Methods

2.1 Notation

- Typography: Sets will be written using the blackboard bold typography (e.g., $\mathbb{L}$, $\mathbb{W}$, and $\mathbb{Q}$), unless a prior convention has already assigned it another symbol. Matrices will be in bold uppercase (e.g., $\mathbf{P}$ and $\mathbf{M}$), tuples, vectors, and multi-vectors will be in bold lowercase (e.g., $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{g}$), and most other constructions (e.g., scalars and functions) will have plain typography (e.g., a, A). The unit pseudo-scalar (of geometric algebra), imaginary number, and identity matrix will be $\mathbf{i}$, i, and $\mathbf{I}$, respectively.

- Sets: The projection of a tuple $\mathbf{p}$ will be $\text{proj}_i(\mathbf{p})$. As an example, the elements of $\mathbb{R}^2 = \mathbb{R}_1 \times \mathbb{R}_2$ are denoted as $\mathbf{p} = (x, y)$. The projection operators are $\text{proj}_1(\mathbf{p}) = x$ and $\text{proj}_2(\mathbf{p}) = y$. If projected over a set, the results are $\text{proj}_1(\mathbb{R}^2) = \mathbb{R}_1$ and $\text{proj}_2(\mathbb{R}^2) = \mathbb{R}_2$. The size of a set $\mathbb{X}$ is $|\mathbb{X}|$.

  The symbol $\cong$ indicates a group isomorphism relation between two sets. The symbol $\simeq$ indicates equality if defined, or both undefined otherwise.

- Analysis: The asterisk $z^\dagger$ denotes the complex conjugate of z.

- Matrix: The Dirac gamma matrices are γ_0, γ_1, γ_2, and γ_3. The Pauli matrices are σ_x, σ_y, and σ_z. The dagger $\mathbf{M}^\dagger$ denotes the conjugate transpose of $\mathbf{M}$. The commutator is defined as $[\mathbf{M}, \mathbf{P}] : \mathbf{MP} - \mathbf{PM}$ and the anti-commutator is defined as $\{\mathbf{M}, \mathbf{P}\} : \mathbf{MP} + \mathbf{PM}$.

- Geometric algebra: The elements of an arbitrary curvilinear geometric basis will be denoted as $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (such that $\mathbf{e}_\nu \cdot \mathbf{e}_\mu = g_{\mu\nu}$), and $\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_n$ (such that $\hat{\mathbf{x}}_\mu \cdot \hat{\mathbf{x}}_\nu = \eta_{\mu\nu}$) if they are orthonormal. A geometric algebra of m dimensions over a field $\mathbb{F}$ is denoted as $\mathbb{G}(m, \mathbb{F})$. The grades of a multi-vector is denoted as $\langle \mathbf{v} \rangle_k$. Specifically, $\langle \mathbf{v} \rangle_0$ is a scalar, $\langle \mathbf{v} \rangle_1$ is a vector, $\langle \mathbf{v} \rangle_2$ is a bi-vector, $\langle \mathbf{v} \rangle_{n-1}$ is a pseudo-vector, and $\langle \mathbf{v} \rangle_n$ is a pseudo-scalar. A scalar and a vector such as $\langle \mathbf{v} \rangle_0 + \langle \mathbf{v} \rangle_1$ form a para-vector, and a combination of even grades ($\langle \mathbf{v} \rangle_0 + \langle \mathbf{v} \rangle_2 + \langle \mathbf{v} \rangle_4 + \ldots$) or odd grades ($\langle \mathbf{v} \rangle_1 + \langle \mathbf{v} \rangle_3 + \ldots$) form even or odd multi-vectors, respectively.

  Let $\mathbb{G}(2, \mathbb{R})$ be the 2D geometric algebra over the reals. We can write a general multi-vector of $\mathbb{G}(2, \mathbb{R})$ as $\mathbf{u} = a + \mathbf{x} + \mathbf{b}$, where a is a scalar, $\mathbf{x}$ is a vector, and $\mathbf{b}$ is a pseudo-scalar.

  Let $\mathbb{G}(4, \mathbb{R})$ be the 4D geometric algebra over the reals. We can write a general multi-vector of $\mathbb{G}(4, \mathbb{R})$ as $\mathbf{u} = a + \mathbf{x} + \mathbf{f} + \mathbf{v} + \mathbf{b}$, where a is a scalar, $\mathbf{x}$ is a vector, $\mathbf{f}$ is a bivector, $\mathbf{v}$ is a pseudo-vector, and $\mathbf{b}$ is a pseudo-scalar.

2.2 Geometric constraints

Definition 1 (Geometric constraints). *Let $\mathbf{M}$ be a $n \times n$ matrix and let $\mathbb{Q}$ be a statistical ensemble. Then, this equality constraint is*

$$\operatorname{tr}\overline{\mathbf{M}} = \sum_{q\in\mathbb{Q}} \rho(q)\operatorname{tr}\mathbf{M}(q), \tag{10}$$

which is called a geometric constraint.

The geometric constraint can also be represented using a multi-vector $\mathbf{u}$ *of a geometric algebra* $\mathbb{G}(4,\mathbb{R})$

$$\operatorname{tr}\overline{\mathbf{u}} = \sum_{q\in\mathbb{Q}} \rho(q)\operatorname{tr}\mathbf{u}(q), \tag{11}$$

The trace $\operatorname{tr}\overline{\mathbf{M}}$ *or* $\operatorname{tr}\overline{\mathbf{u}}$ *denotes the expectation value of the statistically weighted sum of matrices* $\mathbf{M}(q)$ *or of multi-vectors* $\mathbf{u}(q)$ *parametrized over the ensemble* $\mathbb{Q}$*.*

2.3 Geometric representation of matrices

The notation will be significantly improved if we use a geometric representation of matrices, which we introduce in this section.

2.3.1 Geometric representation of 2x2 real matrices

Let $\mathbb{G}(2,\mathbb{R})$ be the 2D geometric algebra over the reals. We can write a general multi-vector of $\mathbb{G}(2,\mathbb{R})$ as

$$\mathbf{u} = a + \mathbf{x} + \mathbf{b}, \tag{12}$$

where, a is a scalar, $\mathbf{x}$ is a vector, and $\mathbf{b}$ is a pseudo-scalar.

Each multi-vector has a structure-preserving (addition/multiplication) matrix representation.

Definition 2 (2D geometric representation)**.**

$$a + x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + b\hat{\mathbf{x}}\wedge\hat{\mathbf{y}} \cong \begin{bmatrix} a+x & -b+y \\ b+y & a-x \end{bmatrix} \tag{13}$$

The converse is also true; each 2×2 real matrix is represented as a multi-vector of $\mathbb{G}(2,\mathbb{R})$.

We can define the determinant using constructs of geometric algebra[3]. The determinant of $\mathbf{u}$ is

Definition 3 (Geometric representation of the determinant 2D)**.**

$$\begin{aligned} \det \quad : \quad & \mathbb{G}(2,\mathbb{R}) \longrightarrow \mathbb{R} \\ & \mathbf{u} \longmapsto \mathbf{u}^{\ddagger}\mathbf{u}, \end{aligned} \tag{14}$$

where, $\mathbf{u}^{\ddagger}$ is

Definition 4 (Clifford conjugate 2D)**.**

$$\mathbf{u}^{\ddagger} := \langle \mathbf{u} \rangle_0 - \langle \mathbf{u} \rangle_1 - \langle \mathbf{u} \rangle_2. \tag{15}$$

For example,

$$\begin{aligned} \det \mathbf{u} &= (a - \mathbf{x} - \mathbf{b})(a + \mathbf{x} + \mathbf{b}) && (16)\\ &= a^2 - x^2 - y^2 + b^2 && (17)\\ &= \det \begin{bmatrix} a + x & -b + y \\ b + y & a - x \end{bmatrix} && (18) \end{aligned}$$

Finally, we defined the Clifford transpose.

Definition 5 (2D Clifford transpose)**.** *The Clifford transpose is the geometric analogue to the conjugate transpose. The conjugate transpose can be interpreted as a transpose followed by an element-by-element application of the complex conjugate. Here, the Clifford transpose is a transpose followed by an element-by-element application of the Clifford conjugate.*

$$\begin{bmatrix} \mathbf{u}_{00} & \dots & \mathbf{u}_{0n} \\ \vdots & \ddots & \vdots \\ \mathbf{u}_{m0} & \dots & \mathbf{u}_{mn} \end{bmatrix}^{\ddagger} = \begin{bmatrix} \mathbf{u}_{00}^{\ddagger} & \dots & \mathbf{u}_{m0}^{\ddagger} \\ \vdots & \ddots & \vdots \\ \mathbf{u}_{m0} & \dots & \mathbf{u}_{nm}^{\ddagger} \end{bmatrix} \tag{19}$$

If applied to a vector, then

$$\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_m \end{bmatrix}^{\ddagger} = \begin{bmatrix} \mathbf{v}_1^{\ddagger} & \dots \mathbf{v}_m^{\ddagger} \end{bmatrix} \tag{20}$$

2.3.2 Geometric representation of 4x4 real matrices

Let $\mathbb{G}(4, \mathbb{R})$ be the 2D geometric algebra over the reals. We can write a general multi-vector of $\mathbb{G}(4, \mathbb{R})$ as

$$\mathbf{u} = a + \mathbf{x} + \mathbf{f} + \mathbf{v} + \mathbf{b}, \tag{21}$$

where, a is a scalar, $\mathbf{x}$ is a vector, $\mathbf{f}$ is a bi-vector, $\mathbf{v}$ is a pseudo-vector, and $\mathbf{b}$ is a pseudo-scalar.

Each multi-vector has a structure-preserving (addition/multiplication) matrix representation. The multi-vectors of $\mathbb{G}(4, \mathbb{R})$ are represented as follows:

Definition 6 (4D geometric representation)**.**

$$
\begin{aligned}
&a + t\gamma_0 + x\gamma_1 + y\gamma_2 + z\gamma_3 \\
&\quad + f_{01}\gamma_0 \wedge \gamma_1 + f_{02}\gamma_0 \wedge \gamma_2 + f_{03}\gamma_0 \wedge \gamma_3 + f_{23}\gamma_2 \wedge \gamma_3 + f_{13}\gamma_1 \wedge \gamma_3 + f_{12}\gamma_1 \wedge \gamma_2 \\
&\quad + v_t\gamma_1 \wedge \gamma_2 \wedge \gamma_3 + v_x\gamma_0 \wedge \gamma_2 \wedge \gamma_3 + v_y\gamma_0 \wedge \gamma_1 \wedge \gamma_3 + v_z\gamma_0 \wedge \gamma_1 \wedge \gamma_2 \\
&\quad + b\gamma_0 \wedge \gamma_1 \wedge \gamma_2 \wedge \gamma_3
\end{aligned}
$$

$$
\cong \begin{bmatrix} a + x_0 - if_{12} - iv_3 & f_{13} - if_{23} + v_2 - iv_1 & -ib + x_3 + f_{03} - iv_0 & x_1 - ix_2 + f_{01} - if_{02} \\ -f_{13} - if_{23} - v_2 - iv_1 & a + x_0 + if_{12} + iv_3 & x_1 + ix_2 + f_{01} + if_{02} & -ib - x_3 - f_{03} - iv_0 \\ -ib - x_3 + f_{03} + iv_0 & -x_1 + ix_2 + f_{01} - if_{02} & a - x_0 - if_{12} + iv_3 & f_{13} - if_{23} - v_2 + iv_1 \\ -x_1 - ix_2 + f_{01} + if_{02} & -ib + x_3 - f_{03} + iv_0 & -f_{13} - if_{23} + v_2 + iv_1 & a - x_0 + if_{12} - iv_3 \end{bmatrix} \tag{22}
$$

Here, the converse is not true, that is, it is only a subset of a 4×4 real matrix that can be represented as a multi-vector of $\mathbb{G}(4, \mathbb{R})$. However, the 4D multi-vector only grabs a fraction of 4×4 complex matrices. Moreover, since both the 4×4 matrices and multi-vectors of $\mathbb{G}(4, \mathbb{R})$ have 16 independent variables and their determinants are real-valued, they have similar properties.

In 4D, we can define the determinant solely using constructs of geometric algebra[3]. The determinant of $\mathbf{u}$ is

Definition 7 (4D geometric representation of determinant)**.**

$$
\begin{aligned}
\det \quad : \qquad & \mathbb{G}(4, \mathbb{R}) \longrightarrow \mathbb{R} && (23) \\
& \mathbf{u} \longmapsto \lfloor \mathbf{u}^\ddagger \mathbf{u} \rfloor_{3,4} \mathbf{u}^\ddagger \mathbf{u}, && (24)
\end{aligned}
$$

where, $\mathbf{u}^\ddagger$ is

Definition 8 (4D Clifford conjugate)**.**

$$
\mathbf{u}^\ddagger := \langle \mathbf{u} \rangle_0 - \langle \mathbf{u} \rangle_1 - \langle \mathbf{u} \rangle_2 + \langle \mathbf{u} \rangle_3 + \langle \mathbf{u} \rangle_4, \tag{25}
$$

where $\lfloor \mathbf{m} \rfloor_{\{3,4\}}$ is the blade-conjugate of degrees 3 and 4 (flipping the plus sign to a minus sign for blades 3 and 4)

$$
\lfloor \mathbf{u} \rfloor_{\{3,4\}} := \langle \mathbf{u} \rangle_0 + \langle \mathbf{u} \rangle_1 + \langle \mathbf{u} \rangle_2 - \langle \mathbf{u} \rangle_3 - \langle \mathbf{u} \rangle_4. \tag{26}
$$

2.4 Unitary gauge (Recap)

Quantum electrodynamics are obtained by gauging the wave function with $U(1)$. The $U(1)$ invariance results from the usage of the complex norm in ordinary quantum theory. A parametrization of ψ over a differentiable manifold is required to support this derivation. Localizing the invariance group $\theta \to \theta(x)$ over the said parametrization yields the corresponding covariant derivative, which is given by

$$
D_\mu = \partial_\mu + iqA_\mu(x), \tag{27}
$$

where, $A_\mu(x)$ is the gauge field.
If a gauge transformation is applied to ψ and A_μ, then

$$\psi \to e^{-iq\theta(x)}\psi \quad \text{and} \quad A_\mu \to A_\mu + \partial_\mu \theta(x). \tag{28}$$

The covariant derivative is

$$\begin{aligned} D_\mu \psi &= \partial_\mu \psi + iqA_\mu \psi && (29) \\ &\to \partial_\mu (e^{-iq\theta(x)}\psi) + iq(A_\mu + \partial_\mu \theta(x))(e^{-iq\theta(x)}\psi) && (30) \\ &= e^{-iq\theta(x)} D_\mu \psi. && (31) \end{aligned}$$

Finally, the field is expressed as

$$F_{\mu\nu} = [\mathcal{D}_\mu, \mathcal{D}_\nu], \tag{32}$$

where $\mathcal{D}_\mu$ is the covariant derivative with respect to the potential one-form $A_\mu = A_\mu^\alpha T_\alpha$, and T_α are the generators of the lie algebra of $U(1)$.

3 Result

3.1 Non-relativistic quantum mechanics

In this section, we recover non-relativistic quantum mechanics using the Lagrange multipliers method and a geometric constraint.

Instead of the Boltzmann entropy, we use the Shannon entropy.

$$S = -\sum_{q \in \mathbb{Q}} \rho(q) \ln \rho(q) \tag{33}$$

In statistical mechanics, we use "scalar" constraints on the entropy, such as the energy meter and volume meter. These are sufficient for recovering the Gibbs ensemble but insufficient for recovering quantum mechanics. A "specialized" geometric constraint which is invariant for a complex phase, is defined as

$$\operatorname{tr}\begin{bmatrix} 0 & -\overline{b} \\ \overline{b} & 0 \end{bmatrix} = \sum_{q \in \mathbb{Q}} \rho(q) \operatorname{tr}\begin{bmatrix} 0 & -b(q) \\ b(q) & 0, \end{bmatrix} \tag{34}$$

where, $\begin{bmatrix} a(q) & -b(q) \\ b(q) & a(q) \end{bmatrix} \cong a(q) + ib(q)$ is the matrix representation of the complex numbers. Similar to the energy meter or volume meter, geometric instruments produce a sequence of measurements converging to an expectation value, but such measurements have a phase invariance. The trace grants and enforces this phase invariance.

The Lagrangian equation that maximizes the entropy subject to this constraint is

$$\mathcal{L} = -\sum_{q\in\mathbb{Q}} \rho(q)\ln(q) + \alpha\left(1 - \sum_{q\in\mathbb{Q}}\rho(q)\right) + \tau\left(\operatorname{tr}\begin{bmatrix}0 & -\overline{b}\\ \overline{b} & 0\end{bmatrix} - \sum_{q\in\mathbb{Q}}\rho(q)\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{35}$$

Maximizing this equation for ρ by posing $\frac{\partial\mathcal{L}}{\partial\rho(q)} = 0$, we obtain

$$\frac{\partial\mathcal{L}}{\partial\rho(q)} = -\ln\rho(q) - 1 - \alpha - \tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{36}$$

$$0 = \ln\rho(q) + 1 + \alpha + \tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{37}$$

$$\implies \ln\rho(q) = -1 - \alpha - \tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{38}$$

$$\implies \rho(q) = \exp(-1-\alpha)\exp\left(-\tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{39}$$

$$= \frac{1}{Z(\tau)}\det\exp\left(-\tau\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right), \tag{40}$$

where, $Z(\tau)$ is obtained as

$$1 = \sum_{q\in\mathbb{Q}}\exp(-1-\alpha)\exp\left(-\tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{41}$$

$$\implies (\exp(-1-\alpha))^{-1} = \sum_{q\in\mathbb{Q}}\exp\left(-\tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{42}$$

$$Z(\tau) := \sum_{q\in\mathbb{Q}}\det\exp\left(-\tau\begin{bmatrix}0 & -b(q)\\ b(q) & 0.\end{bmatrix}\right) \tag{43}$$

The exponential of the trace is equal to the determinant of the exponential via the relation $\det\exp A \equiv \exp\operatorname{tr} A$.

Finally, we obtained

$$\rho(\tau, q) = \frac{1}{Z(\tau)}\det\exp\left(-\tau\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{44}$$

$$\cong |\exp -i\tau b(q)|^2 \qquad \text{Born rule} \tag{45}$$

Renaming $\tau \to t/\hbar$ and $b(q) \to H(q)$ recovers the familiar form of

$$\rho(q) = \frac{1}{Z}\left|\exp(-itH(q)/\hbar)\right|^2. \tag{46}$$

or in even a more familiar form

$$\rho(q) = \frac{1}{Z} |\psi(q)|^2 , \text{ where } \psi(q) = \exp(-itH(q)/\hbar). \tag{47}$$

With this, we can show that all three Dirac Von-Neumann axioms and the Born rule are satisfied, thus providing an origin story for quantum mechanics linked to entropy and geometry.

Indeed, from (47), we can identify the wave function as the vector of some orthogonal space (in this case, a complex Hilbert space) and the partition function as its inner product expressed as

$$Z = \langle\psi|\psi\rangle . \tag{48}$$

After normalization, the physical states are its unit vectors. The probability of any particular state is given as

$$\rho(q) = \frac{1}{\langle\psi|\psi\rangle} (\psi(q))^\dagger \psi(q). \tag{49}$$

Finally, any self-adjoint matrix, defined as $\langle \mathbf{O}\psi|\psi\rangle = \langle\psi|\mathbf{O}\psi\rangle$, will correspond to a real-valued statistical mechanics observable if measured in its eigenbasis.

The equivalence is complete.

3.2 Probability measure of all geometric measurements

Here, we investigate the arbitrary geometric constraint

$$\operatorname{tr} \overline{\mathbf{M}} = \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{M}(q), \tag{50}$$

where $\mathbf{M}$ is the arbitrary $n \times n$ matrix.

We note that we could have used an arbitrary multi-vector $\mathbf{u}$ of $\mathbb{G}(4, \mathbb{R})$ instead of $\mathbf{M}$; the steps of the derivation are the same.

The Lagrange equation used to maximize the entropy subject to this constraint is expressed as

$$\mathcal{L} = -\sum_{q\in\mathbb{Q}} \rho(q) \ln(q) + \alpha \left(1 - \sum_{q\in\mathbb{Q}} \rho(q)\right) + \tau \left(\operatorname{tr} \overline{\mathbf{M}} - \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{M}(q)\right), \tag{51}$$

where α and τ are the Lagrange multipliers.

Maximizing this equation for ρ by posing $\frac{\partial \mathcal{L}}{\partial \rho(q)} = 0$, we obtain

$$\frac{\partial \mathcal{L}}{\partial \rho(q)} = -\ln \rho(q) - 1 - \alpha - \tau \operatorname{tr} \mathbf{M}(q) \tag{52}$$

$$0 = \ln \rho(q) + 1 + \alpha + \tau \operatorname{tr} \mathbf{M}(q) \tag{53}$$

$$\implies \ln \rho(q) = -1 - \alpha - \tau \operatorname{tr} \mathbf{M}(q) \tag{54}$$

$$\implies \rho(q) = \exp(-1-\alpha)\exp(-\tau \operatorname{tr} \mathbf{M}(q)) \tag{55}$$

$$= \frac{1}{Z(\tau)} \det \exp(-\tau \mathbf{M}(q)) \tag{56}$$

where, $Z(\tau)$ is obtained as

$$1 = \sum_{q \in \mathbb{Q}} \exp(-1-\alpha)\exp(-\tau \operatorname{tr} \mathbf{M}(q)) \tag{57}$$

$$\implies (\exp(-1-\alpha))^{-1} = \sum_{q \in \mathbb{Q}} \exp(-\tau \operatorname{tr} \mathbf{M}(q)) \tag{58}$$

$$Z(\tau) \coloneqq \sum_{q \in \mathbb{Q}} \det \exp(-\tau \mathbf{M}(q)) \tag{59}$$

The resulting probability measure is

$$\rho(q, \tau) = \frac{1}{Z(\tau)} \det \exp(-\tau \mathbf{M}(q)), \tag{60}$$

where

$$Z(\tau) = \sum_{q \in \mathbb{Q}} \det \exp(-\tau \mathbf{M}(q)). \tag{61}$$

Posing $\psi(q, \tau) = \exp(-\tau \mathbf{M}(q))$, we can write $\rho(q, \tau) = \det \psi(q, \tau)$, where the determinant acts as a "generalized Born rule," connecting in this case a general linear amplitude to a real number representing a probability.

It is the sophistication of the general linear amplitude along with the determinant acting as a "generalized Born rule" that increases the opportunity to support both general relativity and the standard model, while nonetheless behaving as a consistent physical system due to having its origins solidly anchored in the robust framework of statistical mechanics.

4 Geometric foundation of physics

In this section, we investigate the main result as a general linear quantum theory. In addition, we introduce the *algebra of geometric observables* applicable to the general linear wave function. The 2D case constitutes a special case whose definitions have direct correspondences with those of ordinary quantum mechanics. The 4D case is significantly more sophisticated than the 2D case, and will be investigated immediately after.

4.1 2D axiomatic definition of the algebra

Let $\mathbb{V}$ be an m-dimensional vector space over $\mathbb{G}(2,\mathbb{R})$. A subset of vectors in $\mathbb{V}$ forms an algebra of observables $\mathcal{A}(\mathbb{V})$ if the following holds:

A) $\forall \boldsymbol{\psi} \in \mathcal{A}(\mathbb{V})$, the sesquilinear map

$$\begin{aligned} \langle \cdot, \cdot \rangle \quad : \qquad \mathbb{V} \times \mathbb{V} &\longrightarrow \mathbb{G}(2,\mathbb{R}) \\ \langle \mathbf{u}, \mathbf{v} \rangle &\longmapsto \mathbf{u}^{\ddagger}\mathbf{v} \end{aligned} \tag{62}$$

is positive-definite when $\mathbf{u} = \mathbf{v}$, that is $\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle > 0$

B) $\forall \boldsymbol{\psi} \in \mathcal{A}(\mathbb{V})$. Then, for each element $\psi(q) \in \boldsymbol{\psi}$, the function

$$\rho(\psi(q), \boldsymbol{\psi}) = \frac{1}{\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle} \psi(q)^{\ddagger} \psi(q) \tag{63}$$

is positive-definite: $\rho(\psi(q), \boldsymbol{\psi}) > 0$

We note the following comments and definitions:

- From A) and B), it follows that $\forall \boldsymbol{\psi} \in \mathcal{A}(\mathbb{V})$, the probabilities sum up to unity:

$$\sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q), \boldsymbol{\psi}) = 1 \tag{64}$$

- $\boldsymbol{\psi}$ is called a *natural* (or physical) state.
- $\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle$ is called the *partition function* of $\boldsymbol{\psi}$.
- If $\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle = 1$, then $\boldsymbol{\psi}$ is called a unit vector.
- $\rho(q, \boldsymbol{\psi})$ is called the *probability measure* (or generalized Born rule) of $\psi(q)$.
- The set of all matrices $\mathbf{T}$ acting on $\boldsymbol{\psi}$ as $\mathbf{T}\boldsymbol{\psi} \to \boldsymbol{\psi}'$, making the sum of probabilities normalized (invariant).

$$\sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q), \mathbf{T}\boldsymbol{\psi}) = \sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q), \boldsymbol{\psi}) = 1 \tag{65}$$

are the *natural* transformations of $\boldsymbol{\psi}$.

- A matrix $\mathbf{O}$ such that $\forall \mathbf{u} \forall \mathbf{v} \in \mathcal{A}(\mathbb{V})$:

$$\langle \mathbf{O}\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{O}\mathbf{v} \rangle \tag{66}$$

is called an observable.

- The expectation value of an observable $\mathbf{O}$ is

$$\langle \mathbf{O} \rangle = \frac{1}{\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle} \langle \mathbf{O}\boldsymbol{\psi}, \boldsymbol{\psi} \rangle \tag{67}$$

4.2 Observable in 2D — self-adjoint operator

The general case of an observable in 2D is investigated in this section. A matrix $\mathbf{O}$ is an observable if it is a self-adjoint operator. It is defined as

$$\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle = \langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle \tag{68}$$

$\forall \mathbf{u} \forall \mathbf{v} \in \mathbb{V}$.

Setup: Let $\mathbf{O} = \begin{bmatrix} \mathbf{o}_{00} & \mathbf{o}_{01} \\ \mathbf{o}_{10} & \mathbf{o}_{11} \end{bmatrix}$ be an observable. Let $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ be two two-state vectors of multi-vectors $\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{bmatrix}$ and $\boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\psi}_1 \\ \boldsymbol{\psi}_2 \end{bmatrix}$. Here, the components $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, $\boldsymbol{\psi}_1$, $\boldsymbol{\psi}_2$, $\mathbf{o}_{00}$, $\mathbf{o}_{01}$, $\mathbf{o}_{10}$, $\mathbf{o}_{11}$ are multi-vectors of $\mathbb{G}(2, \mathbb{R})$.

Derivation: 1. Calculate $\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle$:

$$\begin{aligned} 2\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle &= (\mathbf{o}_{00}\boldsymbol{\phi}_1 + \mathbf{o}_{01}\boldsymbol{\phi}_2)^\ddagger \boldsymbol{\psi}_1 + \boldsymbol{\psi}_1^\ddagger (\mathbf{o}_{00}\boldsymbol{\phi}_1 + \mathbf{o}_{01}\boldsymbol{\phi}_2) \\ &\quad + (\mathbf{o}_{10}\boldsymbol{\phi}_1 + \mathbf{o}_{11}\boldsymbol{\phi}_2)^\ddagger \boldsymbol{\psi}_2 + \boldsymbol{\psi}_2^\ddagger (\mathbf{o}_{10}\boldsymbol{\phi}_1 + \mathbf{o}_{11}\boldsymbol{\phi}_2) \qquad (69) \\ &= \boldsymbol{\phi}_{\mathbf{1}}{}^\ddagger \mathbf{o}_{00}^\ddagger \boldsymbol{\psi}_1 + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{01}^\ddagger \boldsymbol{\psi}_1 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{00} \boldsymbol{\phi}_1 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{01} \boldsymbol{\phi}_2 \\ &\quad + \boldsymbol{\phi}_1^\ddagger \mathbf{o}_{10}^\ddagger \boldsymbol{\psi}_2 + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{11}^\ddagger \boldsymbol{\psi}_2 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{10} \boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{11} \boldsymbol{\phi}_2 \qquad (70) \end{aligned}$$

2. Now, $\langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle$:

$$\begin{aligned} 2\langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle &= \boldsymbol{\phi}_1^\ddagger (\mathbf{o}_{00}\boldsymbol{\psi}_1 + \mathbf{o}_{01}\boldsymbol{\psi}_2) + (\mathbf{o}_{00}\boldsymbol{\psi}_1 + \mathbf{o}_{01}\boldsymbol{\psi}_2)^\ddagger \boldsymbol{\phi}_1 \\ &\quad + \boldsymbol{\phi}_2^\ddagger (\mathbf{o}_{10}\boldsymbol{\psi}_1 + \mathbf{o}_{11}\boldsymbol{\psi}_2) + (\mathbf{o}_{10}\boldsymbol{\psi}_1 + \mathbf{o}_{11}\boldsymbol{\psi}_2)^\ddagger \boldsymbol{\phi}_1 \qquad (71) \\ &= \boldsymbol{\phi}_1^\ddagger \mathbf{o}_{00} \boldsymbol{\psi}_1 + \boldsymbol{\phi}_1^\ddagger \mathbf{o}_{01} \boldsymbol{\psi}_2 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{00}^\ddagger \boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{01}^\ddagger \boldsymbol{\phi}_1 \\ &\quad + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{10} \boldsymbol{\psi}_1 + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{11} \boldsymbol{\psi}_2 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{10}^\ddagger \boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{11}^\ddagger \boldsymbol{\phi}_1 \qquad (72) \end{aligned}$$

For $\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle = \langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle$ to be realized, these relations must hold:

$$\mathbf{o}_{00}^\ddagger = \mathbf{o}_{00} \tag{73}$$

$$\mathbf{o}_{01}^\ddagger = \mathbf{o}_{10} \tag{74}$$

$$\mathbf{o}_{10}^\ddagger = \mathbf{o}_{01} \tag{75}$$

$$\mathbf{o}_{11}^\ddagger = \mathbf{o}_{11}. \tag{76}$$

Therefore, $\mathbf{O}$ must be equal to its own Clifford transpose. Thus, $\mathbf{O}$ is an observable iff

$$\mathbf{O}^\ddagger = \mathbf{O}, \tag{77}$$

which is equivalent to the self-adjoint operator $\mathbf{O}^\dagger = \mathbf{O}$ of complex Hilbert spaces.

The geometric sophistication of this geometric observable allows the probability measure to retain invariance over a larger class of geometric transformations than what is possible with unitary transformation. These transformations are sufficiently flexible to support gravity while retaining valid observable statistics.

4.3 Observable in 2D — eigenvalues / spectral theorem

The application of the spectral theorem to $\mathbf{O}^\ddagger = \mathbf{O}$ such that its eigenvalues are real is as follows: Consider

$$\mathbf{O} = \begin{bmatrix} a_{00} & a - x\mathbf{e}_1 - y\mathbf{e}_2 - b\mathbf{e}_{12} \\ a + x\mathbf{e}_1 + y\mathbf{e}_2 + b\mathbf{e}_{12} & a_{11} \end{bmatrix}, \tag{78}$$

It follows that $\mathbf{O}^\ddagger = \mathbf{O}$

$$\mathbf{O}^\ddagger = \begin{bmatrix} a_{00} & a - x\mathbf{e}_1 - y\mathbf{e}_2 - b\mathbf{e}_{12} \\ a + x\mathbf{e}_1 + y\mathbf{e}_2 + b\mathbf{e}_{12} & a_{11} \end{bmatrix}, \tag{79}$$

This example is the most general 2×2 matrix $\mathbf{O}$ such that $\mathbf{O}^\ddagger = \mathbf{O}$.

The eigenvalues are obtained as

$$0 = \det(\mathbf{O} - \lambda \mathbf{I}) = \det \begin{bmatrix} a_{00} - \lambda & a - x\mathbf{e}_1 - y\mathbf{e}_2 - b\mathbf{e}_{12} \\ a + x\mathbf{e}_1 + y\mathbf{e}_2 + b\mathbf{e}_{12} & a_{11} - \lambda \end{bmatrix}, \tag{80}$$

This implies that

$$0 = (a_{00} - \lambda)(a_{11} - \lambda) - (a - x\mathbf{e}_1 - y\mathbf{e}_2 - b\mathbf{e}_{12})(a + x\mathbf{e}_1 + y\mathbf{e}_2 + b\mathbf{e}_{12} + a_{11}) \tag{81}$$

$$0 = (a_{00} - \lambda)(a_{11} - \lambda) - (a^2 - x^2 - y^2 + b^2), \tag{82}$$

Finally,

$$\lambda = \{\frac{1}{2}\left(a_{00} + a_{11} - \sqrt{(a_{00} - a_{11})^2 + 4(a^2 - x^2 - y^2 + b^2)}\right), \tag{83}$$

$$\frac{1}{2}\left(a_{00} + a_{11} + \sqrt{(a_{00} - a_{11})^2 + 4(a^2 - x^2 - y^2 + b^2)}\right)\} \tag{84}$$

Note that, in the case where $a_{00} - a_{11} = 0$, the roots would be complex if $a^2 - x^2 - y^2 + b^2 < 0$, but we already stated that the determinant of real matrices must be greater than zero because of the exponential maps to the orientation-preserving general linear group. Therefore, it is the case where $a^2 - x^2 - y^2 + b^2 > 0$ because this expression is the determinant of the multi-vector. Consequently, for orientation-preserving transformations, $\mathbf{O}^\ddagger = \mathbf{O}$ implies that its roots are real-valued, thus constituting a “geometric” observable in the traditional sense of an observable whose eigenvalues are real-valued.

4.4 2D left action

A left action on the wave function $\mathbf{T}\left|\psi\right\rangle$ connects to the bilinear form as $\left\langle\psi\right|\mathbf{T}^\ddagger\mathbf{T}\left|\psi\right\rangle$. The invariance requirement on $\mathbf{T}$ is

$$\left\langle\psi\right|\mathbf{T}^\ddagger\mathbf{T}\left|\psi\right\rangle = \left\langle\psi|\psi\right\rangle. \tag{85}$$

Therefore, we are interested in the group of matrices such that

$$\mathbf{T}^\ddagger\mathbf{T} = I. \tag{86}$$

Let us consider a two-state system. A general transformation is

$$\mathbf{T} = \begin{bmatrix} u & v \\ w & x \end{bmatrix}, \tag{87}$$

where u, v, w, x are 2D multi-vectors. The expression $\mathbf{T}^\ddagger\mathbf{T}$ is

$$\mathbf{T}^\ddagger\mathbf{T} = \begin{bmatrix} v^\ddagger & u^\ddagger \\ w^\ddagger & x^\ddagger \end{bmatrix}\begin{bmatrix} v & w \\ u & x \end{bmatrix} = \begin{bmatrix} v^\ddagger v + u^\ddagger u & v^\ddagger w + u^\ddagger x \\ w^\ddagger v + x^\ddagger u & w^\ddagger w + x^\ddagger x \end{bmatrix} \tag{88}$$

For the results to be the identity, it must be the case where

$$v^\ddagger v + u^\ddagger u = 1 \tag{89}$$
$$v^\ddagger w + u^\ddagger x = 0 \tag{90}$$
$$w^\ddagger v + x^\ddagger u = 0 \tag{91}$$
$$w^\ddagger w + x^\ddagger x = 1 \tag{92}$$

This is the case if

$$\mathbf{T} = \frac{1}{\sqrt{v^\ddagger v + u^\ddagger u}}\begin{bmatrix} v & u \\ -e^\varphi u^\ddagger & e^\varphi v^\ddagger \end{bmatrix}, \tag{93}$$

where u, v are 2D multi-vectors and e^φ is a unit multi-vector. Comparatively, the unitary case is obtained when the vector part of the multi-vector vanishes $\mathbf{x} \to 0$, and is

$$\mathbf{U} = \frac{1}{\sqrt{|a|^2 + |b|^2}}\begin{bmatrix} a & b \\ -e^{i\theta} b^\dagger & e^{i\theta} a^\dagger \end{bmatrix}. \tag{94}$$

We can show that $\mathbf{T}^\ddagger\mathbf{T} = I$ as follows:

$$\implies \mathbf{T}^{\ddagger}\mathbf{T} = \frac{1}{v^{\ddagger}v + u^{\ddagger}u} \begin{bmatrix} v^{\ddagger} & -e^{-\varphi}u \\ u^{\ddagger} & e^{-\varphi}v \end{bmatrix} \begin{bmatrix} v & u \\ -e^{\varphi}u^{\ddagger} & e^{\varphi}v^{\ddagger} \end{bmatrix} \tag{95}$$

$$= \frac{1}{v^{\ddagger}v + u^{\ddagger}u} \begin{bmatrix} v^{\ddagger}v + u^{\ddagger}u & v^{\ddagger}u - v^{\ddagger}u \\ u^{\ddagger}v - u^{\ddagger}v & u^{\ddagger}u + v^{\ddagger}v \end{bmatrix} \tag{96}$$

$$= I. \tag{97}$$

In the case where $\mathbf{T}$ and $|\psi\rangle$ are n-dimensional, we can find an expression for it starting from a diagonal matrix.

$$\mathbf{D} = \begin{bmatrix} e^{x_1\hat{\mathbf{x}}+y_1\hat{\mathbf{y}}+ib_1} & 0 \\ 0 & e^{x_2\hat{\mathbf{x}}+y_2\hat{\mathbf{y}}+ib_2} \end{bmatrix}, \tag{98}$$

where, $\mathbf{T} = P\mathbf{D}P^{-1}$. It follows easily that $D^{\ddagger}D = I$ because each diagonal entry produces unity: $e^{-x_1\hat{\mathbf{x}}-y_1\hat{\mathbf{y}}-ib_1}e^{x_1\hat{\mathbf{x}}+y_1\hat{\mathbf{y}}+ib_1} = 1$.

An arbitrary matrix $\mathbf{T}$ such that $\mathbf{T}^{\ddagger}\mathbf{T} = I$ can be expressed as an exponential

$$\mathbf{T} = \exp(-\tau\mathbf{A}), \tag{99}$$

where $\mathbf{A}^{\ddagger} = -\mathbf{A}$. Then,

$$\exp(-\tau\mathbf{A})^{\ddagger}\exp(-\tau\mathbf{A}) = \exp(\tau\mathbf{A})\exp(-\tau\mathbf{A}) = I \tag{100}$$

An example of a matrix $\mathbf{A}$ is

$$\begin{bmatrix} \mathbf{x}_1 + \mathbf{b}_1 & \mathbf{x}_3 + \mathbf{b}_3 & \dots \\ \mathbf{x}_3 + \mathbf{b}_3 & \mathbf{x}_2 + \mathbf{b}_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{101}$$

In ordinary quantum mechanics, the equivalent relation is $(e^{iH})^{\dagger}e^{iH} = e^{-iH}e^{iH} = I$

4.5 Dynamics in 2D

We will now derive the relativistic dynamics in 2D.

We start with this equation

$$\exp(-\delta\tau\mathbf{A})\,|\psi(\tau)\rangle = |\psi(\tau+\delta\tau)\rangle\,. \tag{102}$$

Now we approximate the exponential into a power series

$$\exp(-\delta\tau\mathbf{A})\,|\psi(\tau)\rangle \approx 1 - \delta\tau\mathbf{A}\,|\psi(\tau)\rangle\,. \tag{103}$$

We continue as follows

$$(1 - \delta\tau \mathbf{A}) |\psi(\tau)\rangle = |\psi(\tau + \delta\tau)\rangle \tag{104}$$
$$|\psi(\tau)\rangle - \delta\tau \mathbf{A} |\psi(\tau)\rangle = |\psi(\tau + \delta\tau)\rangle \tag{105}$$
$$-\delta\tau \mathbf{A} |\psi(\tau)\rangle = |\psi(\tau + \delta\tau)\rangle - |\psi(\tau)\rangle \tag{106}$$
$$-\mathbf{A} |\psi(\tau)\rangle = \frac{|\psi(\tau + \delta\tau)\rangle - |\psi(\tau)\rangle}{\delta\tau} \tag{107}$$
$$-\mathbf{A} |\psi(\tau)\rangle = \frac{d |\psi(\tau)\rangle}{d\tau}. \tag{108}$$

In the case where we pose $\mathbf{x} \to 0$ (this corresponds to a reduction of the SL(2, $\mathbb{R}$) to the SO(1,1)), then $\mathbf{A}$ reduces to a matrix of pseudo-scalars, which can be written as $\mathbf{A}_{\mathbf{x}\to 0} = \mathbf{iB}$. The corresponding equation is:

$$-\mathbf{iB} |\psi(\tau)\rangle = \frac{d |\psi(\tau)\rangle}{d\tau}, \tag{109}$$

This compares to the Schrödinger equation which is

$$-i\mathbf{H} |\psi(\tau)\rangle = \frac{d |\psi(\tau)\rangle}{d\tau}, \tag{110}$$

The wave function is the solution to this differential equation and is given as

$$\psi(\tau) = \exp(-\tau \mathbf{iB} + a) \tag{111}$$

However, despite being nearly identical to the Schrödinger, here our equation Lorentz is invariant due to the pseudo-scalar being a geometric object — we can see it as follows:

$$\psi^\ddagger(\tau)\hat{\mathbf{x}}_0\psi(\tau) = \exp(-\tau\mathbf{iB} + a)^\ddagger \hat{\mathbf{x}}_0 \exp(-\tau\mathbf{iB} + a) \tag{112}$$
$$= \exp(\tau\mathbf{iB} + a)\hat{\mathbf{x}}_0 \exp(-\tau\mathbf{iB} + a) \tag{113}$$
$$= \exp(2a)\exp(\tau\mathbf{iB})\hat{\mathbf{x}}_0 \exp(-\tau\mathbf{iB}) \tag{114}$$
$$= \rho \exp(\tau\mathbf{iB})\hat{\mathbf{x}}_0 \exp(-\tau\mathbf{iB}) \tag{115}$$

But since $\mathbf{i} = \hat{\mathbf{x}}_0\hat{\mathbf{x}}_1$ then $\mathbf{B}$ is bi-vector of $\mathbb{G}(4, \mathbb{R})$ and these corresponds to a Lorentz rotor SO(1,1).

$$\psi^\ddagger(\tau)\hat{\mathbf{x}}_0\psi(\tau) = \rho \exp(\tau\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1\mathbf{B})\hat{\mathbf{x}}_0 \exp(-\tau\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1\mathbf{B}) \tag{116}$$

The expression $\exp(\tau\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1\mathbf{B})\hat{\mathbf{x}}_0\exp(-\tau\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1\mathbf{B})$ maps $\hat{\mathbf{x}}_0$ to a curvilinear basis $\mathbf{e}_0$ via the application of the rotor and its reverse: $\exp(\tau\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1\mathbf{B}) = R(\tau)$ and $\exp(-\tau\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1\mathbf{B}) = \widetilde{R}(\tau)$

$$R(\tau)\hat{\mathbf{x}}_0\widetilde{R}(\tau) = \mathbf{e}_0(\tau) \tag{117}$$

Therefore

$$\psi^{\ddagger}(\tau)\hat{\mathbf{x}}_0\psi(\tau) = \rho\mathbf{e}_0(\tau) \tag{118}$$

In the David Hestenes formulation of the relativistic wave function this is simply the Dirac current, where $\mathbf{e}_0(\tau)$ is interpreted as the velocity v_0, and ρv_0 is the weighted probability that the particle has the given velocity.

In 1+1 spacetime, the other component of the current vector is

$$\psi^{\ddagger}(\tau)\hat{\mathbf{x}}_1\psi(\tau) = \rho\mathbf{e}_1(\tau) \tag{119}$$

David Hestenes[4] shows that this formulation in 4D is equivalent to other formulations for the relativistic wave-function.

4.6 Algebra of geometric observables in 4D

The general case for a vector space over 4×4 matrices is considered.

In 2D, we extended the complex Hilbert space to a "geometric Hilbert space" and found that the familiar properties of the complex Hilbert spaces were transferable to the geometry of the general linear group.

In 4D, a similar correspondence exists but it is less direct.

The culprit in 4D, is that we need four multiplicands $\lfloor\psi^{\ddagger}\psi\rfloor_{3,4}\psi^{\ddagger}\psi$, compared to the 2D case whose determinant is expressible with two multiplicands $\psi^{\ddagger}\psi$, which can be interpreted as an inner product of two vectors. As such, we are unable to produce a sesquilinear form of the inner product as we did for the 2D case. Since there is no satisfactory inner product, there is no Hilbert space in the usual sense of a complete *inner product* space.

Nevertheless, the quantum mechanical "features" (linear transformations, observables as matrix or operators, and interference patterns in the probability measure) remain in the 4D case.

Our aim is to find the space that supports the general linear wave function in 4D.

A 4 degree "inner product" extension to Hilbert spaces can be created to accommodate our structure. To construct it, the role of the inner product is adopted by a degree 4 "inner product" linking four vectors to an element of $\mathbb{G}(4,\mathbb{R})$. In this construction, the typical concepts of quantum mechanics have equivalences, and the sophistication of the degree 4 "inner product" allows the wave function to accommodate all transformations which we believe may be required to support the *complete* quantum mechanical theory in 4D while retaining valid probabilities for its observables.

Let $\mathbb{V}$ be a m-dimensional vector space over the 4×4 real matrices. A subset of vectors in $\mathbb{V}$ forms an algebra of observables $\mathcal{A}(\mathbb{V})$ if the following holds:

1. $\forall \boldsymbol{\phi} \in \mathcal{A}(\mathbb{V})$, the quadri-sesquilinear form

$$\begin{aligned} \langle \cdot, \cdot, \cdot, \cdot \rangle \quad : \qquad & \mathbb{V} \times \mathbb{V} \times \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{G}(4, \mathbb{R}) \\ & \langle \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \rangle \longmapsto \sum_{i=1}^{m} \lfloor u_i^\ddagger v_i \rfloor_{3,4} w_i^\ddagger x_i \end{aligned} \tag{120}$$

   is positive-definite when $\mathbf{u} = \mathbf{v} = \mathbf{w} = \mathbf{x}$; that is $\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle > 0$

2. $\forall \boldsymbol{\phi} \in \mathcal{A}(\mathbb{V})$, then for each element $\psi(q) \in \boldsymbol{\phi}$, the function

$$\rho(\psi(q), \boldsymbol{\phi}) = \frac{1}{\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle} \det \phi(q), \tag{121}$$

   is positive-definite: $\rho(\phi(q), \boldsymbol{\phi}) > 0$

We note the following properties, features, and comments:

- From A) and B), it follows that, $\forall \boldsymbol{\phi} \in \mathcal{A}(\mathbb{V})$, and the probabilities sum to unity.

$$\sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q), \boldsymbol{\phi}) = 1 \tag{122}$$

- $\boldsymbol{\phi}$ is called a *natural* (or physical) state.
- $\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle$ is called the *partition function* of $\boldsymbol{\phi}$.
- If $\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle = 1$, then $\boldsymbol{\phi}$ is called a unit vector.
- $\rho(\phi(q), \boldsymbol{\phi})$ is called the *probability measure* (or generalized Born rule) of $\phi(q)$.
- The set of all matrices $\mathbf{T}$ acting on $\boldsymbol{\phi}$ such as $\mathbf{T}\boldsymbol{\phi} \to \boldsymbol{\phi}'$ makes the sum of probabilities normalized (invariant):

$$\sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q), \mathbf{T}\boldsymbol{\phi}) = \sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q), \boldsymbol{\phi}) = 1 \tag{123}$$

  are the *natural* transformations of $\boldsymbol{\phi}$.

- A matrix $\mathbf{O}$ such that $\forall \mathbf{u} \forall \mathbf{v} \forall \mathbf{w} \forall \mathbf{x} \in \mathbb{V}$:

$$\langle \mathbf{O}\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{O}\mathbf{v}, \mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{v}, \mathbf{O}\mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{O}\mathbf{x} \rangle \tag{124}$$

  is called an observable.

- The expectation value of an observable $\mathbf{O}$ is

$$\langle \mathbf{O} \rangle = \frac{\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle}{\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle} \tag{125}$$

4.6.1 Wave-function

Let us now make a few comments.

In the David Hestenes' notation[4], the wave-function is given as

$$\psi = \sqrt{\rho e^{ib}} R, \tag{126}$$

where ρ represents a scalar probability density ρ, where e^{ib} is a complex phase and where R is a rotor expressed as the exponential of a bi-vector.

In our framework, the 4D probability measure is given from a degree 4 "inner-product" versus the degree 2 inner product of a Hilbert space. To recover the David Hestenes' formulation of the wave function, we must square our wave function (we will also eliminate the terms $\mathbf{x} \to 0$ and $\mathbf{v} \to 0$ from it to reduce the general linear group to spinors).

$$\psi = \phi^2|_{\mathbf{x}\to 0, \mathbf{v}\to 0} = e^{2a+2\mathbf{b}+2\mathbf{f}} = \sqrt{\rho e^{ib}} R \tag{127}$$

In a loose sense, our wave function ϕ can be interpreted, within the context of the probability measure, as the "square" or "double-copy" of an ordinary spinor wave-function. The additional geometric richness of the "double-copy" is required to support the general linear group quantum mechanically. Finally, the probability measure reduces to the familiar wave-function when the general linear group is reduced to spinors, merely by squaring.

Working with the more sophisticated degree 4 "inner product" increases the opportunity to support more of modern physics under a singular framework, because this measure now supports all possible geometric measurement in nature. To realize this potential, we will now investigate observables, gravity, and suggest falsifiable predictions.

4.6.2 Observables

In 4D, an observable must satisfy equation 125:

$$\lfloor (\mathbf{O}\psi)^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi = \lfloor \psi^\ddagger \mathbf{O}\psi \rfloor_{3,4} \psi^\ddagger \psi = \lfloor \psi^\ddagger \psi \rfloor_{3,4} (\mathbf{O}\psi)^\ddagger \psi = \lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \mathbf{O}\psi \tag{128}$$

$$\lfloor \psi^\ddagger \mathbf{O}^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi = \lfloor \psi^\ddagger \mathbf{O}\psi \rfloor_{3,4} \psi^\ddagger \psi = \lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \mathbf{O}^\ddagger \psi = \lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \mathbf{O}\psi \tag{129}$$

Since the middle terms cancel $\lfloor \psi \rfloor_{3,4} \psi^\ddagger = 1$, the relations can be simplified as

$$e^{2a} \lfloor \psi^\ddagger \mathbf{O}^\ddagger \rfloor_{3,4} \psi = e^{2a} \lfloor \psi^\ddagger \mathbf{O} \rfloor_{3,4} \psi = e^{2a} \lfloor \psi^\ddagger \rfloor_{3,4} \mathbf{O}^\ddagger \psi = e^{2a} \lfloor \psi^\ddagger \rfloor_{3,4} \mathbf{O}\psi \tag{130}$$

It follows that an observable must satisfy

$$\lfloor \mathbf{O}^\ddagger \rfloor_{3,4} = \lfloor \mathbf{O} \rfloor_{3,4} = \mathbf{O}^\ddagger = \mathbf{O}. \tag{131}$$

This is readily satisfied in two cases: complex and bi-vector cases.

1. In the first case, if $\mathbf{O} \in \mathbb{C}^{n\times n}$, then the relations are satisfied if $\mathbf{O}$ is self-adjoint $\mathbf{O}^\dagger = \mathbf{O}$. The corresponding invariance group of the evolution of this observable is unitary $U^\dagger U = I$.
2. In the second case, if $\mathbf{O}$ is a bi-vector, it is satisfied if $\mathbf{O}^\ddagger = \mathbf{O}$. The corresponding invariance group of the evolution of this observable is $F^\ddagger F = I$.

As we will now see, if we then demand that each of these two cases, the evolution preserve the invariance of the Dirac current, then the first and second cases correspond to the SU(2) $\times$ U(1) and SU(3) groups, respectively.

4.6.3 SU(2)xU(1) group

We will now investigate the first case that satisfies the 4D relation for the observables. This corresponds to the case where the observables are self-adjoint $\mathbf{O}^\dagger = \mathbf{O}$ and where the evolution is unitary $U^\dagger U = I$. We will be looking for the most general unitary transformation, expressed as a multi-vector of $\mathbb{G}(4, \mathbb{R})$ which leaves the Dirac current invariant.

Let $\mathbf{u} = a + \mathbf{x} + \mathbf{f} + \mathbf{v} + \mathbf{b}$ be an arbitrary multi-vector of $\mathbb{G}(4, \mathbb{R})$, let $\mathbf{M}$ be its matrix representation, and let ψ be the wave-function.

We will now restrict the set of multi-vectors $e^{\mathbf{u}}$ to those multi-vectors that realize the Dirac current and make it remain invariant after transformation. Specifically, we wish to satisfy this relation

$$\psi^\ddagger \gamma_0 \psi = (e^{\mathbf{u}}\psi)^\ddagger \gamma_0 (e^{\mathbf{u}}\psi) \tag{132}$$

Let us now investigate.

Notably, $\mathbf{x}$ and $\mathbf{v}$ anti-commute with γ_0, and therefore must be equal to 0 as they would otherwise not cancel out. Furthermore, the bi-vectors of $\mathbf{u}$ have basis $\gamma_0\gamma_1, \gamma_0\gamma_2, \gamma_0\gamma_3, \gamma_1\gamma_2, \gamma_1\gamma_3$, and $\gamma_2\gamma_3$. Among these, only $\gamma_1\gamma_2, \gamma_1\gamma_3$, and $\gamma_2\gamma_3$ commute with γ_0; therefore, the rest must be equal to 0. Finally, the pseudo-scalar anti-commutes with γ_0, but this is fine as it must cancel in the Dirac current. Therefore, the most general multi-vector that realizes the definition of the Dirac current and retain its invariance is

$$\mathbf{u} \to a + F_{12}\gamma_1\gamma_2 + F_{13}\gamma_1\gamma_3 + F_{23}\gamma_2\gamma_3 + b\gamma_0\gamma_1\gamma_2\gamma_3 \tag{133}$$

To see its physical significance, it suffices to note that $\gamma_1\gamma_2 = I\sigma_3$, $\gamma_1\gamma_3 = I\sigma_2$ and $\gamma_2\gamma_3 = I\sigma_1$. The resulting multi-vector is unitary and is equal to

$$U = e^{\mathbf{u}} = e^{\frac{1}{2}I(F_{23}\sigma_1 + F_{13}\sigma_2 + F_{12}\sigma_3 + b)}. \tag{134}$$

The terms $F_{23}\sigma_1 + F_{13}\sigma_2 + F_{12}\sigma_3$ and b are responsible for the $SU(2)$ and $U(1)$ symmetries, respectively. The detailing of this identification is available in the reference we cite [5, 6], where David Hestenes and later Lasenby constructs the electroweak sector (and discuss the chromodynamics sector) using the geometric algebra associated with such invariance conditions.

4.6.4 SU(3) group

The second case will be investigated in this section. It corresponds to where the observables is given as $\mathbf{O}^{\ddagger} = \mathbf{O}$ and where the evolution is $F^{\ddagger}F = I$.

Let $\mathbf{f}$ be a bi-vector:

$$\mathbf{f} = F_{01}\gamma_0\gamma_1 + F_{02}\gamma_2\gamma_0 + F_{03}\gamma_0\gamma_3 + F_{23}\gamma_2\gamma_3 + F_{13}\gamma_1\gamma_3 + F_{12}\gamma_1\gamma_2. \tag{135}$$

Alternatively, we can write $\mathbf{f}$ as

$$\mathbf{f} = (F_{01} + \mathbf{i}F_{23})\gamma_0\gamma_1 + (F_{02} + \mathbf{i}F_{13})\gamma_2\gamma_0 + (F_{03} + \mathbf{i}F_{12})\gamma_0\gamma_3, \tag{136}$$

where $\mathbf{i}$ is the $\mathbb{G}(4, \mathbb{R})$ pseudo-scalar.

The current $F^{\ddagger}\gamma_0 F$ is

$$F^{\ddagger}\gamma_0 F = -F\gamma_0 F = (F_{01}^2 + F_{02}^2 + F_{03}^2 + F_{13}^2 + F_{23}^2 + F_{12}^2)\gamma_0 \tag{137}$$
$$+ (-2F_{02}F_{12} + 2F_{03}F_{13})\gamma_1 \tag{138}$$
$$+ (-2F_{01}F_{12} + 2F_{03}F_{23})\gamma_2 \tag{139}$$
$$+ (-2F_{01}F_{13} + 2F_{02}E_{23})\gamma_3 \tag{140}$$

For $F^{\ddagger}\gamma_0 F$ to be make the Dirac current retain its invariance $(F\psi)^{\ddagger}\gamma_0 F\psi = \psi^{\ddagger}\gamma_0\psi$, the cross-product must vanish leaving only

$$F^{\ddagger}\gamma_0 F = (F_{01}^2 + F_{02}^2 + F_{03}^2 + F_{13}^2 + F_{23}^2 + F_{12}^2)\gamma_0, \tag{141}$$

which is the SU(3) group.

With the previous SU(2) × U(1) result (case 1) and SU(3) (case 2), the 4D geometric observables produce the symmetry groups associated with modern particle physics, while leaving minimal wiggle room (but probably not exactly 'no room') for anything different.

Here, the SU(2) × U(1) and the SU(3) groups are the result of "casting" the general degree-4 probability measure into a requirement to preservice the invariance of the Dirac current, which is associated with a degree-2 probability. The "casting" reduces the set of all multi-vector transformations $\psi' = \mathbf{u}\psi$ to only those that leave the Dirac current $\psi^{\ddagger}\gamma_0\psi$ invariant. The resulting multi-vectors form the SU(2) × U(1) group in the first satisfiable case of the observable, and the SU(3) group in the second.

4.6.5 Gravity

We have considered many options for gravity including holographic forms of gravity, gravity by quantum entanglement, gravity from entropy (à la Ted Jacobson[7]), gravity by gauging [8, 9, 10, 11], etc.

Of these options, the gauge gravitation theory defined for (up to) the affine gauge, yielding (up to) the metric-affine gravity directly follows from our method and requires no additions or modifications.

In our framework, the general linear gauge symmetry replaces and generalizes the role of the $U(1)$ gauge symmetry in ordinary quantum mechanics. The general linear gauge is present at the level of the probability measure itself; thus, $\mathrm{GL}(n, \mathbb{F})$ is to $\det \psi$ what $U(1)$ is to $\psi^\dagger \psi$. With this gauge, gravity will be the natural motion of all fields and will couple to all Lagrangians constructing consistently with our probability measure.

The generality of the metric-affine gravity exceeds that of general relativity. This generality can be reduced, if needed, to accommodate multiple flavours of gravity; from the Poincaré gauge theory (nonmetricity=0) to the Einstein-Carton variety and finally to standard general relativity (torsion=0). Our strategy is to support the metric-affine theory of gravity in the general case and, only if needed, reduce the extra freedom in the final result.

A particularly interesting special case of the metric-affine gravity is the teleparallel version of general relativity which relies only on the translation group to realize general relativity (along with a special choice of the action). This divorces the translation group from the general linear group; allowing an interpretation of general relativity as the "transporter" (via the translation group) of quantum mechanical information (available via the general linear group wave-function) along the world manifold.

How is the metric-affine theory of gravity realized?

The affine group is the result of supplementing the general linear group with translation via the semidirect product $\mathrm{A}(4, \mathbb{R}) = T(4) \rtimes \mathrm{GL}(4, \mathbb{R})$. Thus, to realize a gauge theory of this group, we have to handle both translations and the general linear group. The general linear group is the default gauge of our probability measure so this should be straightforward, but what about the translations?

So far we have parametrized our wave-function using the elements q of an arbitrary ensemble $\mathbb{Q}$. The first step is to replace $\mathbb{Q}$ with a world manifold $\mathcal{M}$, and the elements q by the points x on the manifold. On such a manifold, the introduction of a parametrization introduces transformational symmetries leading to gauge symmetries.

First, as it is the easier of the two, the general linear group. We now interpret the general linear wave-function as "living" in the tangent space at each point x of the world manifold $\mathcal{M}$. The geometric basis of the multi-vector $(\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ define the tangent space of $\mathcal{M}$.

A general linear transformation is given by

$$\psi'(x) \to g\psi(x)g^{-1}, \tag{142}$$

The determinant will leave the probability measure of the wave function invariant because

$$\det\bigl(g\psi(x)g^{-1}\bigr) = \det\psi(x). \tag{143}$$

The gauge-covariant derivative associated with this transformation is

$$D_\mu\psi = \partial_\mu\psi - [iqA_\mu, \psi]. \tag{144}$$

Finally, the field is given as

$$R_{\mu\nu} = [D_\mu, D_\nu], \tag{145}$$

where $R_{\mu\nu}$ represents the curvature and allows the definition of the Riemann tensor.

We now must support the second gauge which are the translations. The procedure we will use is standard in the literature and so we only provide a brief sketch here. The best primer we have found is detailed in the following reference[11].

To support affine transformations, we enrich the tangent space T_xM at each point x of $\mathcal{M}$ by another point o_x; this creates a tangent affine space A_xM whose elements are $p_x = (o_x, \mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Translations act on o_x and the general linear group acts on $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. We now want to transform a point p_x from the tangent affine space A_xM to a point $p_{\tilde{x}}$ in $A_{\tilde{x}}M$. A translation of a point in A_xM to a point in $A_{\tilde{x}}M$ involves the use of a connection. Since we can transform any point in A_xM to any point $A_{\tilde{x}}M$, there is a gauge symmetry. Finally, to connect $p_{\tilde{x}}$ in $A_{\tilde{x}}M$ to its corresponding point $\tilde{x}$ in $\mathcal{M}$, a soldering form in employed. The end product is that parallel transport within the tangent affine spaces on different points on the manifold corresponds to diffeomorphism at the level of the manifold. This is the origin of gravity within the gauge-theoretical setup.

In the usual metric-affine theory of gravitation, translations corresponds to torsion T, and the general linear group to curvature R (and non-metricity Q). In this interpretation, the general linear wave-function is intimately connected to the curvature (and non-metricity).

In the teleparallel theory of gravity, translations are sufficient for the theory to be equivalent to general relativity. In this case, the translations are divorced from the general linear group, allowing an interpretation of general relativity as the "transporter" of general linear quantum mechanical information along the world manifold. We find this interpretation interesting, but further exploration is of course required.

In this manuscript, our goal is to discuss our framework in the most general sense possible. We feel it is still too early to make a drastic interpretational choice (teleparallel vs metric-affine) at this stage and without further exploration.

5 Step towards falsifiable predictions

A number of falsifiable predictions is listed below.

The main idea is that a general linear wave function would allow a larger class of interference patterns, compared to complex interference. The general linear interference pattern includes all the ways in which space-time can interfere with itself, including those resulting from rotations, boosts, shear, torsion, etc.

It is plausible that an Aharonov–Bohm effect experiment on gravity[12] could detect the general linear phase and patterns identified in this section.

An interference pattern follows from a linear combination of $\mathbf{u}$ and $\mathbf{v}$, and the application of the determinant:

$$\det(\mathbf{u}+\mathbf{v}) = \det\mathbf{u} + \det\mathbf{v} + \text{extra-terms} \tag{146}$$

The sum of the probability is $(\det\mathbf{u}+\det\mathbf{v})$ and the 'extra terms' represents the interference term.

We use the extra terms to define a bilinear form using the dot product notation.

$$\cdot \quad : \qquad \mathbb{G}(2n,\mathbb{R}) \times \mathbb{G}(2n,\mathbb{R}) \longrightarrow \mathbb{R} \tag{147}$$

$$\mathbf{u}\cdot\mathbf{v} \longmapsto \frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det\mathbf{u} - \det\mathbf{v}) \tag{148}$$

For example, in 2D, we have

$$\mathbf{u} = a_1 + x_1\mathbf{e}_1 + y_1\mathbf{e}_2 + b_1\mathbf{e}_{12} \tag{149}$$

$$\mathbf{v} = a_2 + x_2\mathbf{e}_1 + y_2\mathbf{e}_2 + b_2\mathbf{e}_{12} \tag{150}$$

$$\implies \mathbf{u}\cdot\mathbf{v} = a_1a_2 + b_1b_2 - x_1x_2 - y_1y_2. \tag{151}$$

If $\det\mathbf{u} > 0$ and $\det\mathbf{v} > 0$, then $\mathbf{u}\cdot\mathbf{v}$ is always positive, thereby qualifying as a positive-definite inner product, but no greater than either $\det\mathbf{u}$ or $\det\mathbf{v}$, whichever is greater. Therefore, it also satisfies the conditions of an interference term.

- In 2D, the dot product is equivalent to the form

$$\frac{1}{2}(\det(\mathbf{u}+\mathbf{v})-\det\mathbf{u}-\det\mathbf{v}) = \frac{1}{2}\left((\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v})-\mathbf{u}^\ddagger\mathbf{u}-\mathbf{v}^\ddagger\mathbf{v}\right) \tag{152}$$

$$= \mathbf{u}^\ddagger\mathbf{u}+\mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u}+\mathbf{v}^\ddagger\mathbf{v}-\mathbf{u}^\ddagger\mathbf{u}-\mathbf{v}^\ddagger\mathbf{v} \tag{153}$$

$$= \mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u} \tag{154}$$

- In 4D, it is substantially more complex:

$$\frac{1}{2}(\det(\mathbf{u}+\mathbf{v})-\det\mathbf{u}-\det\mathbf{v}) \tag{155}$$

$$= \frac{1}{2}\left(\lfloor(\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v})\rfloor_{3,4}(\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v})-\lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}-\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\right) \tag{156}$$

$$= \frac{1}{2}\left(\lfloor\mathbf{u}^\ddagger\mathbf{u}+\mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u}+\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}(\mathbf{u}^\ddagger\mathbf{u}+\mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u}+\mathbf{v}^\ddagger\mathbf{v})-\dots\right) \tag{157}$$

$$\begin{aligned}
= {} & \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\\
& +\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\\
& +\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\\
& +\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}-\dots
\end{aligned} \tag{158}$$

$$\begin{aligned}
= {} & \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\\
& +\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\\
& +\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\\
& +\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u}+\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v}+\lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u}
\end{aligned} \tag{159}$$

A simpler version of this interference pattern is possible when the general linear group is reduced.

Complex interference:

In 2D, a reduction of the general linear group to the circle group reduces the interference pattern to a complex interference.

$$|\psi_1+\psi_2|^2 = |\psi_1|^2+|\psi_2|^2+2|\psi_1||\psi_2|\cos(\phi_1-\phi_2) \tag{160}$$

Deep spinor interference:

A reduction to the spinor group reduces the interference pattern to a "deep spinor rotation".

Consider a two-state wave function (we note that $[\mathbf{f}, \mathbf{b}] = 0$)).

$$\psi = \psi_1 + \psi_2 = e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} \tag{161}$$

The geometric interference pattern for a full general linear transformation in 4D is given by

$$\lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi. \tag{162}$$

Starting with the sub-product

$$\begin{aligned}
\psi^\ddagger \psi &= (e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2})(e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2}) && (163)\\
&= e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} \\
&\quad + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2} e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2} e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} && (164)\\
&= e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{2\mathbf{b}_2} + e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) && (165)
\end{aligned}$$

The full product is expressed as

$$\begin{aligned}
\lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi &= \left(e^{2a_1} e^{-2\mathbf{b}_1} + e^{2a_2} e^{-2\mathbf{b}_2} + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})\right) \\
&\quad \times \left(e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{2\mathbf{b}_2} + e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}\right) && (166)\\
&= e^{2a_1} e^{-2\mathbf{b}_1} e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_1} e^{-2\mathbf{b}_1} e^{2a_2} e^{2\mathbf{b}_2} + e^{2a_1} e^{-2\mathbf{b}_1} e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) \\
&\quad + e^{2a_2} e^{-2\mathbf{b}_2} e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{-2\mathbf{b}_2} e^{2a_2} e^{2\mathbf{b}_2} + e^{2a_2} e^{-2\mathbf{b}_2} e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) \\
&\quad + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) e^{2a_1} e^{2\mathbf{b}_1} \\
&\qquad + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) e^{2a_2} e^{2\mathbf{b}_2} \\
&\qquad + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) && (167)\\
&= e^{4a_1} + e^{4a_2} + 2e^{2a_1+2a_2} \cos(2b_1 - 2b_2) && (168)\\
&\quad + e^{a_1+a_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})(&& (169)\\
&\quad\ e^{2a_1} (e^{-\mathbf{b}_1+\mathbf{b}_2} + e^{\mathbf{b}_1-\mathbf{b}_2}) && (170)\\
&\qquad + e^{2a_2} (e^{\mathbf{b}_1-\mathbf{b}_2} + e^{-\mathbf{b}_1+\mathbf{b}_2})) && (171)\\
&\quad + e^{2a_1+2a_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})^2 && (172)\\
&= \underbrace{e^{4a_1} + e^{4a_2}}_{\text{sum}} + \underbrace{2e^{2a_1+2a_2} \cos(2b_1 - 2b_2)}_{\text{complex interference}} \\
&\quad + \underbrace{2e^{a_1+a_2} (e^{2a_1} + e^{2a_2})(e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})(\cos(B_1 - B_2)) + e^{2A_1+2A_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})^2}_{\text{deep spinor interference}} && (173)
\end{aligned}$$

6 Discussion

We have recovered the foundations of quantum mechanics using the tools of statistical mechanics to maximize the entropy, and a geometric constraint. In doing so we have replaced the Boltzmann entropy with the Shannon entropy, and this has an impact on the resulting interpretation.

In contrast to the multiple interpretations of quantum mechanics, the interpretation of statistical mechanics is singular, free of paradoxes and obviously devoid of any measurement problem; remarkably, this will carry over to our interpretation of quantum mechanics.

Definition 9 (Metrological interpretation)**.** *There exist instruments that record sequences of measurements on systems. These measurements are unique up to a geometric phase, and the Born rule (including its geometric generalization to the determinant) is the entropy-maximizing measure constrained by the expectation value of these measurements.*

The Lagrange multiplier method, used to maximize the entropy subject to geometric constraints, is the mathematical backbone to this interpretation.

Let us now discuss the definition of the measuring apparatus entailed by this interpretation.

Integrating formally into physics the notion of an instrument or measuring apparatus has been a long standing difficulty. One of the pitfalls is to attribute too much "detailing" to this instrument (for instance defining the instrument as a macroscopic system which amplifies quantum information), as this increases the risk of capturing only a fraction of all possible instruments in nature. Fractional capture is to be avoided, since the instruments are our only "eyes into nature"; consequently the generality of their definition must be on part with the laws of physics themselves.

Do we have any physical theory which already admits a satisfactory definition of the measuring apparatus?

In statistical mechanics, instruments and their effects on systems are incorporated into the mathematical formalism. For instance, an energy meter or volume meter can produce a sequence of measurements whose average converges towards an expectation value, and this constitutes a constraint on the entropy. But the generality (and generalizability) of this definition to all physical system (including quantum and geometrical) was overlooked. In this study, we have capitalized on this definition and we have extended it appropriately.

The instrument is defined as follows:

Definition 10 (Instrument/Measuring Apparatus)**.** *An instrument, or measuring apparatus, is a device that constrains the entropy to an expectation value; or more precisely, an instrument is described by an equality which constrains the entropy to a given exception value.*

From this, one must resist the temptation to extend this definition to single measurements (instead of expectation values), as this definition is by itself equivalent to quantum mechanics; how do we recover geometry and quantum mechanics?

Nature allows for geometrically richer measurements and instrumentations than what is possible to express with simple "scalar" or "phase-less" instruments. For instance, a ruler, clock, and protractor also admit numerical measurements, but they contain geometric phase invariances such as the Lorentz invariance.

In the metrological interpretation it is not the wave function but the existence of such instruments that is taken as axiomatic. Essentially, the interpretation adopts the belief that the laws of physics are entirely determined by the geometrical richness (invariance) of the instruments that are available in nature.

In this study, we interpreted the trace as the expectation value of the eigenvalues of a matrix transformation times the dimension of the vector space. Maximizing the entropy under the constraint of this expectation value introduces various phase-invariances into the resulting probability measure consistent with the available measuring apparatuses. Specifically, the constraint

$$\operatorname{tr}\begin{bmatrix}0 & -\overline{b}\\ \overline{b} & 0\end{bmatrix} = \sum_{q\in\mathbb{Q}} \operatorname{tr}\rho(q)\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{174}$$

induces a complex phase invariance into the probability measure $\rho(q) = |\exp(-i\tau b(q))|^2$, which gives rise to the Born rule and wave function. Moreover, the constraint

$$\operatorname{tr}\overline{\mathbf{M}} = \sum_{q\in\mathbb{Q}} \operatorname{tr}\rho(q)\mathbf{M}(q) \tag{175}$$

induces a general linear phase invariance in the probability measure $\rho(q) = \det\exp\left(-\tau\mathbf{M}(q)\right)$, giving rise to a probability measure supporting multiple gauges and observables commonly in use in modern physics. In each case, we can interpret the constraint as an instrument acting on the system. In the complex phase, we associated the constraint to an incidence counter measuring a particle or photon. Moreover, in the general linear case, we associated the constraint to a measure that is invariant with respect to all changes of coordinates in the general linear phase, such as measurements of the geometry of space-time.

The complete correspondence between an ordinary system of statistical mechanics and ours is as follows.

Table 1: Correspondence

Concept	Statistical Mechanics	Geometric Constraint (Ours)
Entropy	Boltzmann	Shannon
Measure	Gibbs	Born rule on wave function
Constraint	Energy meter	Phase-invariant instrument
Micro-state	Energy values	Possible measurements
Macro-state	Equation of state	Evolution of the wave function
Experience	Ergodic	Message of measurements

In the correspondence, the usage of the Shannon entropy instead of the Boltzmann entropy changes the experience from ergodic to a message (in the sense of the theory of communication of Claude Shannon[13]) of measurements. The receipt of such a message by say, an observer, carries information; it is interpreted as the registration of a 'click'[14] on a screen or other detecting instrument. Using the Shannon entropy, quantum physics can then be interpreted as the probability measure resulting from maximizing the entropy of a message of geometrically invariant measurements received by an observer.

The probabilistic interpretation of the wave function via the Born rule is inherited from statistical mechanics and results from maximizing the entropy under geometric constraints. The wave function is also entailed; hence, it is not considered axiomatic either. However, it is the receipt of a message of the measurements taken by an instrument along with the geometric constraints on the entropy it entails, that is axiomatic.

The axioms of quantum mechanics are recoverable as theorems from the solution $\frac{\partial \mathcal{L}}{\partial \rho} = 0$ for ρ, where,

$$\mathcal{L} = -\sum_{q\in\mathbb{Q}} \rho(q)\ln\rho(q) + \lambda\left(1 - \sum_{q\in\mathbb{Q}}\rho(q)\right) + \tau\left(\operatorname{tr}\overline{\mathbf{M}} - \sum_{q\in\mathbb{Q}}\rho(q)\operatorname{tr}\mathbf{M}(q)\right). \tag{176}$$

Now, let us discuss the wave-function collapse problem:

Specifically, the mathematical foundation of quantum mechanics contains the following axiom: If the measurement of a quantity $\mathbf{O}$ on ψ gives the result o_n, then the state immediately after measurement is given by the normalized projection of ψ onto the eigensubspace of o_n as

$$\psi \implies \frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}} \tag{177}$$

The measurement-collapse problem is superseded as follows: Before the wave function is derived, measurements are assumed to have already been registered by an instrument and are associated with a geometric constraint, which is axiomatic. Registering new measurements in this case does not mean that a wave function has collapsed, but means that we need to adjust the constraints and derive a new wave function consistent with new measurements. Since the wave function is derived by maximizing the entropy constrained by registered measurements, it never undergoes an update from an uncollapsed state to a collapsed state. The collapse problem is a symptom of attributing an ontology to the wave function; however, the ontology belongs to the instruments and their measurements — not the wave function.

For instance, it is by throwing multiple coins into the air and noting that about half land on head and the other half on tail that we can deduce a corresponding probability measure. Such a probability measure cannot be used to derive the result of the next flip but only its expectation value. Likewise, here it is the expectation value of measurements that are used to derive the wave function. The present derivation of the wave function as a solution to a maximization problem on the entropy under a geometric constraint (themselves representing expectation values) is mathematically consistent with this understanding. The connection to statistical mechanics resets our expectation and understanding of the Born rule to be a probability measure whose domains is that of expectation values and not of singular occurrences of events.

Finally, this formulation is consistent with physics being a purely empirical science. Indeed, as all knowledge of nature comes from the instruments that are constructible in it, posing these instruments to be the axiom of physics (rather than posing as axiom the wave function), then using their definition to derive the wave function, makes the mathematics of physics entirely consistent with it being an empirical science.

The full correspondence is also consistent with the general intuition that *random information*, as by definition it cannot be derived from earlier principles, ought to axiomatic. Ultimately, it is sounder to consider *the message of random measurements* to be the axiomatic foundation of the theory rather than the wave-function (a precise and deterministic mathematical equation). As we have shown the later can be derived from the former, but as the lack of satisfactory mechanism for the wave function collapses suggests — not vice-versa.

6.1 Axioms of Physics

We conjecture that the laws of physics are ultimately reducible to these minimal axioms.

Let q be the elements of a statistical ensemble $\mathbb{Q}$, and let $m(q) \in \mathbb{R}$ be an observable of $\mathbb{Q}$.

Axiom 1 (Observability)**.** *The experience of the observer in nature is defined as the receipt of a message* $\mathbf{m} \in \mathbb{R}^n$ *of measurements, as performed on* n *copies of a physical system* $\mathbb{Q}$*.*

Axiom 2 (Representativeness)**.** *The elements of this message are* <u>*representative*</u> *of* $\mathbb{Q}$*: when* $|\mathbf{m}| \to \infty$*, then* $\overline{m} \in \mathbb{R}$ *(i.e. the average of these measurements converges towards a well-defined expectation value).*

Axiom 3 (Comprehensiveness)**.** *The statistical ensemble is comprehensible: when* $|\mathbf{m}| \to \infty$*, then* $\mathbb{Q}\exists$ *(i.e. when* $|\mathbf{m}| \to \infty$ *then all elements in* $\mathbb{Q}$ *are identified).*

Conjecture 1 (Geometricity)**.** *The geometric constraint is sufficiently sophisticated to represent of all possible measurements in nature (yet sufficiently restrictive to represent only those found in nature):*

$$\operatorname{tr}\overline{\mathbf{M}} = \sum_{q\in\mathbb{Q}} \rho(q)\operatorname{tr}\mathbf{M}(q) \tag{178}$$

where $\operatorname{tr}\mathbf{M}(q) = m(q)$ *is a possible measurement, and where* $\mathbf{M}$ *corresponds to a matrix or multi-vector.*

Theorem 1 (Physics)**.** *Maximizing the entropy of the elements of a message of measurement yields, under the geometric constraint, the model of physics consistent with these measurements:*

$$\mathcal{L} = -\sum_{q\in\mathbb{Q}} \rho(q)\ln\rho(q) + \lambda\left(1 - \sum_{q\in\mathbb{Q}}\rho(q)\right) + \tau\left(\operatorname{tr}\overline{\mathbf{M}} - \sum_{q\in\mathbb{Q}}\rho(q)\operatorname{tr}\mathbf{M}(q)\right). \tag{179}$$

Solving for $\partial\mathcal{L}/\partial\rho = 0$ *implies*

$$\rho(q,\tau) = \frac{1}{Z(\tau)}\det\exp(-\tau\mathbf{M}(q)), \tag{180}$$

where

$$Z(\tau) = \sum_{q\in\mathbb{Q}}\det\exp(-\tau\mathbf{M}(q)). \tag{181}$$

where τ*, the Lagrange multiplier, represents the one-parameter group evolution of, in the general case, the orientation preserving general linear group* $\mathrm{GL}^{+}(\mathrm{n},\mathbb{R})$ *(which corresponds to the structure group of a world manifold when* $\mathbb{Q}$ *is equated to* $\mathcal{M}$*, and along with the wave-function and the standard model gauges, comprises the results presented in this study).*

7 Conclusion

In this study, we proposed a geometric constraint, then we used it to maximize the Shannon entropy. This allowed us to derive a probability measures that supports a richer geometry than what was commonly available, and this substantially extends the opportunity to capture all modern physics within a single framework. To accommodate all possible geometric measurements, the wave function of the general linear group was derived and the Born rule was extended to the determinant. A gauge theory of the affine group emerged following parametrization of the wave function in a world manifold. As we have also seen, "casting" the general linear wave function into the definition of the Dirac current reduces the theory to the SU(2)× U(1) and SU(3) groups for the first and second satisfying cases of the 4D observable, respectively. Finally, an interpretation of quantum mechanics, the metrological interpretation, was proposed; the existence of instruments and the measurements they produce acquire the foundational role, and the wave function is derived as a theorem. The interpretation considers that an observer receives a message (theory of communication/Shannon entropy) of phase-invariant measurements, and that the probability measure which maximizes the information of this message is the wave-function/Born rule.

We also state that the theory is not yet complete; for instance, we have not investigated the interaction picture of this probability measure in the context of gravity, nor have we attempted to quantize it, nor have investigated its renormalization potential.

References

[1] Makoto Yamashita (https://mathoverflow.net/users/9942/makoto yamashita). Geometric interpretation of trace. MathOverflow. URL:https://mathoverflow.net/q/46447 (version: 2016-05-17).

[2] Frederick Reif. *Fundamentals of statistical and thermal physics*. Waveland Press, 2009.

[3] Douglas Lundholm and Lars Svensson. Clifford algebra, geometric algebra, and applications. *arXiv preprint arXiv:0907.5356*, 2009.

[4] David Hestenes. Spacetime physics with geometric algebra. *American Journal of Physics*, 71(7):691–714, 2003.

[5] David Hestenes. Space-time structure of weak and electromagnetic interactions. *Foundations of Physics*, 12(2):153–168, 1982.

[6] Anthony Lasenby. Some recent results for $su(3)$ and octonions within the geometric algebra approach to the fundamental forces of nature. *arXiv preprint arXiv:2202.06733*, 2022.

[7] Ted Jacobson. Thermodynamics of spacetime: the einstein equation of state. *Physical Review Letters*, 75(7):1260, 1995.

[8] Ryoyu Utiyama. Invariant theoretical interpretation of interaction. *Physical Review*, 101(5):1597, 1956.

[9] Tom WB Kibble. Lorentz invariance and the gravitational field. *Journal of mathematical physics*, 2(2):212–221, 1961.

[10] Friedrich W Hehl, J Dermott McCrea, Eckehard W Mielke, and Yuval Ne'eman. Metric-affine gauge theory of gravity: field equations, noether identities, world spinors, and breaking of dilation invariance. *Physics Reports*, 258(1-2):1–171, 1995.

[11] Frank Gronwald. Metric-affine gauge theory of gravity: I. fundamental structure and field equations. *International Journal of Modern Physics D*, 6(03):263–303, 1997.

[12] Chris Overstreet, Peter Asenbaum, Joseph Curti, Minjeong Kim, and Mark A Kasevich. Observation of a gravitational aharonov-bohm effect. *Science*, 375(6577):226–229, 2022.

[13] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[14] John A Wheeler. Information, physics, quantum: The search for links. *Complexity, entropy, and the physics of information*, 8, 1990.