Why It Is Better to Define the Laws of Physics as the Solution to an Optimization Problem, Rather Than Axiomatically

Alexandre Harvey-Tremblay[†]

[†]Independent scientist, aht@protonmail.ch
September 14, 2022

Abstract

Mathematically, the laws of physics are represented with axioms (e.g., the Dirac-Von Neumann axioms, the Wightman axioms, etc.). While axioms in logic are held to be true merely by definition, the laws of physics on the other hand are entailed by lab measurements. The existence of this entailment is suggestive of a more appropriate logical structure than axioms to express the laws of physics. This paper presents this structure and makes the case for its supremacy. Specifically, an optimization problem on the entropy of all possible measurements will be introduced, whose solutions are the laws of physics, thus respecting the entailment from measurements to laws. The solution obtained from our approach not only recovers the Dirac–Von Neumann axioms plus the Born rule, but further improves upon them by automatically restricting the observables to no more than the standard model group symmetry SU(3) x SU(2) x U(1), while simultaneously extending the domain of the probability measure exactly enough to support general relativity in the form of a general linear gauge theory. As the solution to an optimization problem, it is arguable that this tight integration constitutes, in this sense, physics' optimized formulation. Finally, our approach further strengthens the foundation of physics superseding it with the group of all measurements as its new and sole axiomatic foundation, and all "theoretical artefacts" (Born rule, probability amplitude, Hilbert space, observables, etc.) are now promoted from axioms to theorems, thus providing a rigorous account for their previously postulated origin.

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

1 Introduction

In this paper, we propose to understand physics as the solution to an optimization problem on the entropy of all possible measurements. Intuitively, we justify this approach based on the realization that the laws of physics can be justified to be valid based on their compatibility with present lab measurements, whereas axioms are true merely by definition. Here, the group of all measurements will be the axioms, and the laws will be derived as theorems/solutions of an optimization problem of the entropy of said measurements.

By solving the problem, an optimized solution of physical laws is obtained. The solution carries numerous advantages regarding optimality, minimality and elegance which translates into physics as yielding no more than exactly the SU(3) x SU(2) X U(1) group symmetry over the observables of the theory, and exactly no less than general relativity over the evolution of its probability measure. We interpret this tight integration as suggestive of the power and efficiency of adequate mathematical optimization.

Secondary results are also presented such as the true origin of the Born rule and the other axioms of quantum physics, an interpretation of quantum mechanics which appears as correct and exclusionary and how the measurement/collapse problem is deprecated by this setup.

To define the optimization problem rigorously, we first introduce the notion of a geometric constraint. Then we will discuss its interpretation and intended usage as a constraint on the entropy.

The construction of a geometric constraint exploits the connection between geometry and the theory of probability via the trace. The trace of a matrix can be understood 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[1].

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 the matrices $\mathbf{M}(q)$ parameterized over the ensemble $\mathbb{Q}$.

Alternatively (and preferably), we may use the geometric algebra to define the constraint. In this case, it is 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 multivector of the real geometric algebra in n dimensions $\mathbb{G}(n,\mathbb{R})$. Although the constraints can be expressed by both the approaches, the use of multivectors instead of matrices highlights the geometric character of the method. More details on geometric algebra (and the present notation) are provided in the Methods section.

Why this equality?

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

How are constraints of this type typically used?

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 that measures the system energy and produces a series of energy measurements $E_1, E_2, \ldots$, which converge 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 and produces a sequence of measured volumes $V_1, V_2, \ldots$, which also converges to an expectation value $\overline{V}$.

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

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

Using equations (3) and (5), a typical statistical mechanical system 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 the Lagrange multipliers.

Therefore, by 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 method, (3) is replaced with $\operatorname{tr} \overline{\mathbf{M}}$, and a geometric constraint is obtained. Instead of energy or volume meters, we have rulers, clocks, protractors, spin meters, dilation meters, and shear meters.

To connect our proposed method with quantum mechanics, the statistical interpretation of the entropy must be altered. In the modified interpretation, the probability measure quantifies 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 entropy (minus the Boltzmann constant); only the final interpretation is changed (this will be further detailed in the discussion, section 6).

The corresponding Lagrange equation is

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

and this equation 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, we introduce a number of tools using geometric algebra, based on the reported study of Lundholm et al. [3]. Specifically, we introduce the notion of a determinant for multivectors as well as the notions of a Clifford conjugate generalizing the complex conjugate. These tools enable us to express our results geometrically.

In the Results section, we present two solutions of the Lagrange equation. The first is the recovery of the standard nonrelativistic quantum mechanics, 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 multivector.

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

Finally, in the Discussion section, we introduce an interpretation of quantum mechanics consistent with its newly revealed origin, as the measure maximizing the Shannon entropy constrained by geometric measurements, namely the metrological interpretation. In this interpretation, the measurements and the associated constraint on the entropy are considered more fundamental than the wavefunction, which is now entirely derivable. The end product is a theory that deprecates the measurement problem, superseding it with theory of instrumentation, and provides a plausible explanation for the origin of quantum mechanics in nature, thereby, tying it to the geometric measurements that are permissible.

2 Methods

2.1 Notation

- Typography:

  Sets are written using the blackboard bold typography (e.g., $\mathbb{L}$, $\mathbb{W}$, and $\mathbb{Q}$), unless a prior convention assigns it another symbol.

  Matrices are in bold uppercase (e.g., $\mathbf{P}$ and $\mathbf{M}$), tuples, vectors, and multivectors are in bold lowercase (e.g., $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{g}$), and most other constructions (e.g., scalars and functions) have plain typography (e.g., a, A).

  The unit pseudo-scalar (of geometric algebra), imaginary number, and identity matrix are $\mathbf{i}$, i, and $\mathbf{I}$, respectively.

- Sets:

  The projection of a tuple $\mathbf{p}$ is $\text{proj}_i(\mathbf{p})$.

  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, then the corresponding 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 both terms are 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 are 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 multivector are 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 bivector, $\langle \mathbf{v} \rangle_{n-1}$ 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 multivectors, respectively.

  Let $\mathbb{G}(2, \mathbb{R})$ be the 2D geometric algebra over the real set.

  We can formulate a general multivector 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 real set.

  In this case also, a general multivector of $\mathbb{G}(4, \mathbb{R})$ can be formulated 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 an* $n \times n$ *matrix and* $\mathbb{Q}$ *be a statistical ensemble.*

Then, this equality constraint is given by

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

which is called a geometric constraint.

The geometric constraint can also be represented using a multivector $\mathbf{u}$ *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 multivectors* $\mathbf{u}(q)$ *parameterized over the ensemble* $\mathbb{Q}$*.*

2.3 Geometric representation of matrices

2.3.1 Geometric representation of 2×2 real matrices

Let $\mathbb{G}(2,\mathbb{R})$ be the 2D geometric algebra over the real set.

We can write a general multivector 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 multivector 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 multivector of $\mathbb{G}(2,\mathbb{R})$.

We can define the determinant using constructs of geometric algebra[3].

Accordingly, 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,

$$\det \mathbf{u} = (a - \mathbf{x} - \mathbf{b})(a + \mathbf{x} + \mathbf{b}) \tag{16}$$

$$= a^2 - x^2 - y^2 + b^2 \tag{17}$$

$$= \det \begin{bmatrix} a+x & -b+y \\ b+y & a-x \end{bmatrix} \tag{18}$$

Finally, we defined the Clifford transpose.

Definition 5 (2D Clifford transpose). *The Clifford transpose is the geometric analogue to the conjugate transpose, which*

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 real set.

We can write a general multivector 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 bivector, $\mathbf{v}$ is a pseudo-vector, and $\mathbf{b}$ is a pseudo-scalar.

In this case also, each multivector has a structure-preserving (addition/multiplication) matrix representation.

The multivectors of $\mathbb{G}(4, \mathbb{R})$ are represented as follows:

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

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

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

In this case, the converse is not true; that is, only a subset of a 4×4 real matrix can be represented as a multivector of $\mathbb{G}(4, \mathbb{R})$. However, the 4D multivector only includes a fraction of the 4×4 complex matrices. Moreover, the 4×4 matrices as well as the multivectors of $\mathbb{G}(4, \mathbb{R})$ have 16 independent variables and their determinants are real-valued; thus, they have similar group properties.

In 4D also, we can define the determinant solely using the constructs of geometric algebra[3].

The determinant of $\mathbf{u}$ is

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

$$\det \quad : \quad \mathbb{G}(4,\mathbb{R}) \longrightarrow \mathbb{R} \tag{23}$$

$$\mathbf{u} \longmapsto \lfloor \mathbf{u}^{\ddagger}\mathbf{u} \rfloor_{3,4} \mathbf{u}^{\ddagger}\mathbf{u}, \tag{24}$$

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

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

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

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

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

2.4 Unitary gauge (Recap)

In quantum electrodynamics, the wavefunction is gauged with $U(1)$.

The $U(1)$ invariance results from the application of the complex norm in ordinary quantum theory.

A parameterization of ψ over a differentiable manifold is required to support this derivation.

Localizing the invariant group $\theta \to \theta(x)$ over the said parameterization 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

$$D_\mu \psi = \partial_\mu \psi + iqA_\mu \psi \tag{29}$$

$$\to \partial_\mu(e^{-iq\theta(x)}\psi) + iq(A_\mu + \partial_\mu\theta(x))(e^{-iq\theta(x)}\psi) \tag{30}$$

$$= e^{-iq\theta(x)} D_\mu \psi. \tag{31}$$

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 elucidate the recovery of the non-relativistic quantum mechanics using the Lagrange multiplier method and a geometric constraint.

As explained before, the Shannon entropy is applied instead of the Boltzmann entropy to achieve the aforementioned goal.

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

In statistical mechanics, we use "scalar" constraints on the entropy, such as energy and volume meters, which are sufficient for recovering the Gibbs ensemble. However, the application of such scalar constraints is insufficient to recover quantum mechanics.

To overcome this limitation, a "specialized" geometric constraint which is invariant for a complex phase will be used. It 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 or volume meters, geometric instruments produce a sequence of measurements that converge to an expectation value, although such measurements exhibit a phase invariance. This phase invariance originates from the trace.

The Lagrangian equation that maximizes the entropy subject to the above-introduced specialized geometric constraint is

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

This equation is maximized for ρ by imposing the condition $\frac{\partial \mathcal{L}}{\partial \rho(q)} = 0$, and the following results are obtained

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

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

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

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

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

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 according to the relation $\det \exp A \equiv \exp \operatorname{tr} A$.

Finally, we obtain

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

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

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 even a more familiar form of

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

With this, we can show that all the three Dirac Von–Neumann axioms as well as the Born rule are satisfied,

which reveals a possible origin of quantum mechanics linked to entropy and geometry.

Indeed, from (47), we can identify the wavefunction 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 become its unit vectors, and

the probability of any particular state is given by

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

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, thereby completing the equivalence.

3.2 Probability measure of all geometric measurements

Here, we explore 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.

Notably, an arbitrary multivector $\mathbf{u}$ of $\mathbb{G}(4, \mathbb{R})$ can be used, instead of $\mathbf{M}$. In both these cases, the steps of the derivation remain the same.

In this case, the Lagrange equation used to maximize the entropy, under 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.

In this case as well, we maximize this equation for ρ using the criterion $\frac{\partial \mathcal{L}}{\partial \rho(q)} = 0$. This operation results in the following:

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

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

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

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

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

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

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

By defining $\psi(q,\tau) := \exp(-\tau \mathbf{M}(q))$, 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.

The sophistication of the general linear amplitude along with the determinant acting as a "generalized Born rule" provides a platform to support both general relativity and the standard model, while behaving as a consistent physical system because of its origins being solidly anchored in the robust framework of statistical mechanics.

4 Geometric foundation of physics

In this section, the analysis of the main result as a general linear quantum theory is presented. In addition, we introduce the *algebra of geometric observables* applicable to the general linear wavefunction.

The 2D definition of the algebra constitutes a special case reminiscent of the definitions of ordinary quantum mechanics. The 4D case is significantly more sophisticated than the 2D case, and is elucidated immediately after the 2D case analysis.

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 $\mathbf{T}\boldsymbol{\psi} \to \boldsymbol{\psi}'$, such that the sum of probabilities remains normalized.

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

  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 shown in this section. A matrix $\mathbf{O}$ is an observable, if it is a self-adjoint operator, and 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 multivectors $\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{bmatrix}$ 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 multivectors 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, $\langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle$:

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

To realize $\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle = \langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle$, the following relations must hold:

$$
\begin{aligned}
\mathbf{o}_{00}^{\ddagger} &= \mathbf{o}_{00} && (73) \\
\mathbf{o}_{01}^{\ddagger} &= \mathbf{o}_{10} && (74) \\
\mathbf{o}_{10}^{\ddagger} &= \mathbf{o}_{01} && (75) \\
\mathbf{o}_{11}^{\ddagger} &= \mathbf{o}_{11}. && (76)
\end{aligned}
$$

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

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

which is the correspond 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 remain invariant over a class of geometric transformations that is larger than that of the unitary transformations.

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 shown below:

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

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

This implies that

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

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

Finally,

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

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

Notably, in the case where $a_{00} - a_{11} = 0$, the roots would be complex if $a^2 - x^2 - y^2 + b^2 < 0$. However, we already stated that the determinant of real matrices must be greater than zero because of the exponential mapping to the orientation-preserving general linear group. Therefore, in this case, $a^2 - x^2 - y^2 + b^2 > 0$, because this expression is the determinant of the multivector.

Consequently, under the 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 wavefunction $\mathbf{T}|\psi\rangle$ connects to the bilinear form as $\langle\psi|\mathbf{T}^\ddagger\mathbf{T}|\psi\rangle$.

The invariance requirement on $\mathbf{T}$ is

$$\langle\psi|\mathbf{T}^\ddagger\mathbf{T}|\psi\rangle = \langle\psi|\psi\rangle. \quad (85)$$

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

$$\mathbf{T}^\ddagger\mathbf{T} = I. \quad (86)$$

Let us consider a two-state system.

A general transformation of such a system is represented by

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

where u, v, w, x are the 2D multivectors.

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 $\mathbf{T}^{\ddagger}\mathbf{T} = I$, it must be the case that

$$v^{\ddagger}v + u^{\ddagger}u = 1 \tag{89}$$

$$v^{\ddagger}w + u^{\ddagger}x = 0 \tag{90}$$

$$w^{\ddagger}v + x^{\ddagger}u = 0 \tag{91}$$

$$w^{\ddagger}w + x^{\ddagger}x = 1 \tag{92}$$

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 the 2D multivectors, and e^{φ} is a unit multivector.

Comparatively, the unitary case is obtained when the vector part of the multivector vanishes, i.e., $\mathbf{x} \to 0$, and we obtain

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

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 identify its general expression 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 $\mathbf{T}$, which yields $\mathbf{T}^{\ddagger}\mathbf{T} = I$, can be expressed as an exponential as follows:

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

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

In this section, we chalk out the derivation of the relativistic dynamics in 2D, starting with the following 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 as

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

The process is continued as follows

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

When we consider $\mathbf{x} \to 0$ (this corresponds to a reduction of SL(2, $\mathbb{R}$) to SO(1,1)), $\mathbf{A}$ 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 equation is similar to the Schrödinger equation:

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

The wavefunction is the solution to this differential equation and is given by

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

Despite being nearly identical to the Schrödinger equation, our equation is Lorentz invariant, because the pseudo-scalar is a geometric object. This can be explained as follows:

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

Because $\mathbf{i} = \hat{\mathbf{x}}_0\hat{\mathbf{x}}_1$, $\mathbf{B}$ is a bivector of $\mathbb{G}(4,\mathbb{R})$ it 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 relativistic wavefunction formulation put forward by David Hestenes, 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 the 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 the other formulations of the relativistic wavefunction.

4.6 Algebra of geometric observables in 4D

In this section, 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.

Although a similar correspondence exists in 4D, it is less recognizable because in 4D, we need four multiplicands $\lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi$. By contrast, in the 2D case, the determinant can be expressed using only two multiplicands $\psi^\ddagger \psi$, which can be interpreted as an inner product of two vectors.

Thus, in 4D, we cannot produce a sesquilinear form of the inner product, similar to the 2D case, and the absence of a satisfactory inner product indicates that 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 wavefunction in 4D.

A four degree "inner product" extension to the Hilbert space can be created to accommodate our structure.

To construct such a Hilbert space for our structure, a degree-four "inner product" is devised, which performs the role of the inner product, mapping 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-four "inner product" allows the wavefunction 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 Wavefunction

Here, we present a few comments on the wavefunction formulation.

In the David Hestenes' notation[4], the wavefunction is expressed by

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

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

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

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

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

This measure now supports all the possible geometric measurements in nature, and thus, the more sophisticated degree-four "inner product" facilitates the inclusion of a larger chunk of modern physics under a single framework. To highlight the potential of this measure to support all the possible geometric measurements in nature, we examine the observables and gravity as well as suggest falsifiable predictions.

4.6.2 Observables

In 4D, an observable must satisfy the equation 125:

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

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

Because 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, viz. the complex and bivector cases.

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

Under the condition that the evolution in each of these two cases preserve the invariance of the Dirac current, the first and second cases correspond to the SU(2) × U(1) and SU(3) groups, respectively.

4.6.3 $SU(2) \times U(1)$ group

Here, we show 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 multivector 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 multivector of $\mathbb{G}(4,\mathbb{R})$, $\mathbf{M}$ be its matrix representation, and ψ be the wavefunction.

We restrict the set of the multivectors $e^{\mathbf{u}}$ to those that realize the Dirac current and retain its invariance after the 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 bivectors 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 multivector 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 is sufficient 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 multivector 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 details of this identification process is available in [5, 6]. David Hestenes and later Lasenby constructed the electroweak sector (and discussed the chromodynamics sector) using the geometric algebra associated with such invariance conditions.

4.6.4 SU(3) group

In the second case, the observables are given by $\mathbf{O}^\ddagger = \mathbf{O}$, and the evolution is $F^\ddagger F = I$.

Let $\mathbf{f}$ be a bivector:

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

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

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

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) $\times$ U(1) and SU(3) groups are the result of "casting" the general degree-four probability measure into the definition of the Dirac current, i.e., the invariance of the Dirac current should be preserved, which is associated with a degree-two probability. The "casting" reduces the set of all multivector transformations $\psi' = \mathbf{u}\psi$ to only those that leave the Dirac current $\psi^{\ddagger}\gamma_0\psi$ invariant.

The resulting multivectors form the SU(2) $\times$ U(1) group in the first satisfiable case of the observable, and the SU(3) group in the second.

4.6.5 Gravity

We considered numerous 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], and so on.

Among these, 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. Indeed, the probability measure is invariant for the general linear gauge.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 the fields and will couple to all the Lagrangians consistent 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 flavors of gravity, from the Poincaré gauge theory (nonmetricity=0) to the Einstein–Cartan 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 the appropriate physical evidence is eventually obtained, to reduce the extra freedom in the final result.

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?

Thus far, we have parameterized our wavefunction 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 parameterization introduces transformational symmetries, leading to gauge symmetries.

First, let us investigate the general linear group. We interpret the general linear wavefunction as "living" in the tangent space at each point x of the world manifold $\mathcal{M}$. The geometric basis of the multivector $(\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ 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 renders the probability measure of the wavefunction invariant because

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

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 the 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. Because 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 wavefunction is intimately connected to the curvature (and non-metricity).

5 Step toward falsifiable predictions

A number of falsifiable predictions is listed below.

The main idea is that a general linear wavefunction 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, and so on.

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 not 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 wavefunction (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

$$\psi^\ddagger \psi = (e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2})(e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2}) \tag{163}$$

$$\begin{aligned} &= e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} \\ &\quad + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2} e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2} e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} \end{aligned} \tag{164}$$

$$= e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{2\mathbf{b}_2} + e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) \tag{165}$$

The full product is expressed as

$$\begin{aligned} \lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi = &\big(e^{2a_1} e^{-2\mathbf{b}_1} + e^{2a_2} e^{-2\mathbf{b}_2} + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})\big) \\ &\times \big(e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{2\mathbf{b}_2} + e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}\big) \end{aligned} \tag{166}$$

$$\begin{aligned} &= e^{2a_1} e^{-2\mathbf{b}_1} e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_1} e^{-2\mathbf{b}_1} e^{2a_2} e^{2\mathbf{b}_2} + e^{2a_1} e^{-2\mathbf{b}_1} e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) \\ &\quad + e^{2a_2} e^{-2\mathbf{b}_2} e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{-2\mathbf{b}_2} e^{2a_2} e^{2\mathbf{b}_2} + e^{2a_2} e^{-2\mathbf{b}_2} e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) \\ &\quad + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) e^{2a_1} e^{2\mathbf{b}_1} \\ &\quad\quad + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) e^{2a_2} e^{2\mathbf{b}_2} \\ &\quad\quad + e^{a_1+a_2} e^{-\mathbf{b}_1-\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) e^{a_1+a_2} e^{\mathbf{b}_1+\mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) \end{aligned} \tag{167}$$

$$= e^{4a_1} + e^{4a_2} + 2e^{2a_1+2a_2} \cos(2b_1 - 2b_2) \tag{168}$$

$$+ e^{a_1+a_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})(\tag{169}$$

$$e^{2a_1} (e^{-\mathbf{b}_1+\mathbf{b}_2} + e^{\mathbf{b}_1-\mathbf{b}_2}) \tag{170}$$

$$+ e^{2a_2} (e^{\mathbf{b}_1-\mathbf{b}_2} + e^{-\mathbf{b}_1+\mathbf{b}_2})) \tag{171}$$

$$+ e^{2a_1+2a_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})^2 \tag{172}$$

$$\begin{aligned} &= \underbrace{e^{4a_1} + e^{4a_2}}_{\text{sum}} + \underbrace{2e^{2a_1+2a_2} \cos(2b_1 - 2b_2)}_{\text{complex interference}} \\ &\quad + \underbrace{2e^{a_1+a_2} (e^{2a_1} + e^{2a_2})(e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})(\cos(B_1 - B_2)) + e^{2A_1+2A_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1})^2}_{\text{deep spinor interference}} \end{aligned} \tag{173}$$

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, because 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. However, 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 geometrically richer measurements and instrumentations, which are not possible to express with simple "scalar" or "phase-less" instruments. For instance, a ruler, clock, and protractor also admit numerical measurements; however, they contain geometric phase invariances, such as the Lorentz invariance.

In the metrological interpretation, the existence of such instruments, not the wavefunction, 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 wavefunction. 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 used in modern physics; specifically, those of general relativity and the standard model.

In each case, we can interpret the constraint as an instrument acting on the system.

In the complex phase, we associate the constraint to an incidence counter measuring a particle or photon. Moreover, in the general linear case, we associate the constraint to a measure that is invariant with respect to all coordinate changes 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 wavefunction
Constraint	Energy meter	Phase-invariant instrument
Micro-state	Energy values	Possible measurements
Macro-state	Equation of state	Evolution of the wavefunction
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 be interpreted as the probability measure resulting from the maximization of the entropy of a message of geometrically invariant measurements received by an observer.

The probabilistic interpretation of the wavefunction via the Born rule is inherited from statistical mechanics and results from the maximization of the entropy under geometric constraints.

The wavefunction is also entailed, and hence, not considered axiomatic either. Instead, the receipt of a message of the measurements taken by an instrument, along with the geometric constraints on the corresponding entropy, 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 wavefunction 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 the 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 wavefunction 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 wavefunction has collapsed, but implies that we need to adjust the constraints and derive a new wavefunction consistent with new measurements. Because the wavefunction is derived by maximizing the entropy constrained by the registered measurements, it never updates from an uncollapsed state to a collapsed state. The collapse problem is a symptom of attributing an ontology to the wavefunction; however, the ontology belongs to the instruments and their measurements — not the wavefunction.

For instance, we can deduce a probability measure by throwing multiple coins into air and noting that about half of these coins land on head and the other half on tail. Such a probability measure cannot be used to derive the result of the next flip, but only its expectation value. Likewise, here, the expectation value of the measurements is used to derive the wavefunction. The present derivation of the wavefunction 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 domain 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 can be constructed, postulating these instruments to be the axiom of physics (rather than the wavefunction), and then using their definition to derive the wavefunction, 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* ought to be axiomatic, as by definition it cannot be derived from any earlier principles. Ultimately, it is viable to consider *the message of random measurements*, rather than the wavefunction (which is a precise and deterministic mathematical equation), to be the axiomatic foundation of the theory. As shown, the latter can be derived from the former, but not vice versa, which is suggested by the lack of a satisfactory mechanism for the wavefunction collapse in the usual interpretation.

6.1 Axioms of Physics

We propose that the laws of physics are ultimately entailed only and entirely by the following minimal axioms regarding measurements.

Let q be the elements of a statistical ensemble $\mathbb{Q}$ and $\forall q \in \mathbb{Q} : m(q) \in \mathbb{R}$ 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 n measurements performed on n identical copies of $\mathbb{Q}$.*

Axiom 2 (Representativeness)**.** *Observations are <u>representative</u> in the limit: 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)**.** *Observations are <u>comprehensive</u> in the limit: when $|\mathbf{m}| \to \infty$, then $\mathbb{Q}$ is well-defined (i.e., all the elements in $\mathbb{Q}$ are identified).*

Conjecture 1 (Geometricity)**.** *The geometric constraint is sufficiently sophisticated to represent all the possible measurements 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* $\mathbf{M}$ *corresponds to a matrix or multivector.*

Conjecture 2 (Geometric Totality)**.** *The geometric constraint is sufficiently restrictive to represent only the measurements that are possible in nature.*

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 the Lagrange multiplier τ *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 wavefunction and the standard model gauges, comprises the results presented in this study).*

7 Conclusion

In this paper, we proposed a geometric constraint, which was used to maximize the Shannon entropy. This geometric constraint allowed us to derive a probability measure that supports a geometry richer than that available commonly, and this substantially extends the opportunity to capture all the modern physics phenomena within a single framework. To accommodate all the possible geometric measurements, the wavefunction 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 the parameterization of the wavefunction in a world manifold. Evidently, "casting" the general linear wavefunction 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, viz. the metrological interpretation, is proposed; the existence of instruments and the measurements they produce acquire the foundational role, and the wavefunction is derived as a theorem. In this interpretation, it is considered 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 wavefunction plus Born rule.

Other aspects, such as the renormalization potential of the gravitational theory or the interaction picture, may be investigated in a future study.

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.