A Gravitized Standard Model Is Found as the Solution to the Problem of Maximizing the Entropy of All Geometric Measurements

Alexandre Harvey-Tremblay[†]

[†]Independent scientist, aht@protonmail.ch
February 18, 2023

Abstract

In modern theoretical physics, the laws of physics are directly represented with axioms (e.g., the Dirac–Von Neumann axioms, the Wightman axioms, Newton's laws of motion). Although in logic axioms are held to be true merely by definition, in physics the laws are entailed by laboratory measurements. The existence of this entailment suggests a more suitable logical structure than axioms to represent the laws of physics. This paper introduces this logical structure and then demonstrates its supremacy. Specifically, an optimization problem on the entropy of all geometric measurements is introduced. Its solution is an optimized version of the Dirac–Von Neumann axioms that automatically restricts its observables to no more than the standard model group symmetry SU(3) and SU(2) x U(1) while simultaneously extending its probability measure to the theory of general relativity (i.e., it is a "gravitized" standard model). Remarkably, this result only holds in 4 dimensions.

1 Introduction

In modern theoretical physics, physical laws are expressed as axioms (e.g., the Dirac–Von Neumann axioms, the Wightman axioms, Newton's laws of motion). The theorems provable by these axioms are the predictions of the theory. If the predictions are invalidated by laboratory measurements, the postulated laws are deemed falsified and new (and possibly more appropriate) laws are instead postulated.

In this scenario, it is the theorems (predictions) of the theory that are used (along with laboratory experimental data) to invalidate its axioms (laws).

In logic, however, axioms define what is true in a theory, and its theorems cannot, of course, invalidate the axioms they depend on.

Thus, there is a dissimilarity between the use of axioms in physics versus their use in logic.

Since the laws of physics require a more complex interplay between axioms, theorems and their invalidations than merely the unidirectional entailment between axioms and theorems found in logic, the question of the applicability of axioms to express the laws of physics arise.

We believe that axioms are in fact inappropriate as a logical tool to define the laws of physics. We intend to show that correcting the axiomatic entailment between the laws and the measurements yields a significantly superior and optimized formulation of fundamental physics.

In our proposal, laboratory measurements entail the mathematical expression of those measurements and it is this expression, not the laws of physics, that will constitute the axioms of our system. The laws of physics are defined as the solution to a carefully crafted optimization problem on the entropy of all measurements.

The solution to this optimization problem is a novel and optimized formulation of fundamental physics. It yields the SU(3) and the SU(2)×U(1) group symmetry over the observables of the theory of the general relativity. Remarkably, no other solutions are possible and this solution only holds in 3+1 dimensions. We interpret this tight configuration as suggestive of the power and efficiency of defining the laws of physics as the solution to a mathematical optimization problem.

We believe that our optimized formulation is unlikely to have been obtained by trial and error or by traditional methods, making our optimization problem a key step in its derivation.

In essence, it is easier, from laboratory measurements, to "guess" the right mathematical expression for all possible measurements, than it is to "guess" the right laws of physics from those same measurements. The distance one must travel in "guessing space" is much shorter in the former case than in the latter, and this is beneficial.

Secondary results are also presented and follow directly from our solution, such as the mathematical origin of the Born rule, the proof of the axioms of quantum physics, the identification of the correct interpretation of quantum mechanics, and the deprecation of the measurement/collapse problem.

To define the problem in full rigor, we first introduce the key structure that makes our approach possible: the geometric constraint, then we give its rationale.

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 multiplied by the dimension of the vector space, and the eigenvalues as 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 geometric algebra to define the constraint as (the notation is explained in section 2)

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

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 approaches, the use of multivectors instead of matrices highlights the geometric character of the method.

Now, we discuss its rationale.

Constraints are used in statistical mechanics 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's 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

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

where

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

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 protractors, boost meters, dilation meters, and shear meters.

We believe that, by limiting its definition of constraints to scalar expressions, statistical physics missed the opportunity to capture all possible geometric measurements available in nature.

Our geometric constraint represents the set all geometric measurements we believe are possible in nature. Specifically, the constraint will support observing the distortions produced by any geometric transformation of events in nature, and the resulting probability measure will preserve the expectation value of these distortions up to a phase or symmetry group.

Many of us are familiar with the expression: "If all one has is a hammer, everything looks like a nail". This idea is essentially geometrized: If all one has are protractors, boost meters, dilation meters, and shear meters, then everything looks geometrically invariant. For instance, a statistical system measured exclusively with a protractor will carry, following our entropy maximization procedure, the rotation symmetry in the probability measure of the events it measured.

Finally, let us note that we will maximize the Shannon entropy and not the Boltzmann entropy. In our interpretation, the resulting probability measure will quantify the information associated with the receipt of a message of measurements by an observer. Using the Shannon entropy does not change the form of the mathematical equation for entropy (minus the Boltzmann constant); only the final interpretation is changed (further details on the interpretation of quantum mechanics are provided in 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, which is our main result.

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, 4]. Specifically, we utilize 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 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 two-dimensions (2D) and 3+1 dimensions (3+1D), consistent with the general solution. Remarkably, in 3+1D, we obtain a sophisticated relation for the transformation invariance, which together with the wavefunction satisfy the SU(2) $\times$ U(1) and SU(3) gauge symmetries. We also obtain a gravitized quantum theory which incorporates general relativity.

Finally, in the Discussion section, we introduce an interpretation of quantum mechanics consistent with its newly revealed origin. It is the measure maximizing the Shannon entropy constrained by geometric measurements, namely the metrological interpretation. In this interpretation, the measurements and 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, supersedes it with theory of instrumentation, and provides a plausible explanation for the origin of quantum mechanics in nature, tying it entirely to its geometric measurability.

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 an homomorphism.

- 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 $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ (such that $\mathbf{e}_\nu \cdot \mathbf{e}_\mu = g_{\mu\nu}$), and $\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_n$ (such that $\hat{\mathbf{x}}_\mu \cdot \hat{\mathbf{x}}_\nu = \eta_{\mu\nu}$) if they are orthonormal.

  A geometric algebra of $m + n$ dimensions over a field $\mathbb{F}$ is denoted as $\mathcal{G}(\mathbb{F}^{m,n})$.

  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 $(\langle \mathbf{v} \rangle_0 + \langle \mathbf{v} \rangle_2 + \langle \mathbf{v} \rangle_4 + \dots)$ or odd grades $(\langle \mathbf{v} \rangle_1 + \langle \mathbf{v} \rangle_3 + \dots)$ form even or odd multivectors, respectively.

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

  We can formulate a general multivector of $\mathcal{G}(\mathbb{R}^2)$ 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 $\mathcal{G}(\mathbb{R}^{3,1})$ be the 3+1D geometric algebra over the real set.

  Similarly, a general multivector of $\mathcal{G}(\mathbb{R}^{3,1})$ 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.*

The geometric constraint is

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

The geometric constraint can also be represented using a multivector $\mathbf{u}$ *of a geometric algebra* $\mathcal{G}(\mathbb{R}^{m,n})$

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

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 in 2D

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

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

In geometric algebra, the determinant[4] of a multivector $\mathbf{u}$ can be defined as:

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

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

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

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

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

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 define 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 in 3+1D

Let $\mathcal{G}(\mathbb{R}^{3,1})$ be the 3+1D geometric algebra over the real set.

We can write a general multivector of $\mathcal{G}(\mathbb{R}^{3,1})$ 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.

Similarly, each multivector has a structure-preserving (addition/multiplication) matrix representation.

The multivectors of $\mathcal{G}(\mathbb{R}^{3,1})$ 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 complex matrices can be represented as a multivector of $\mathcal{G}(\mathbb{R}^{3,1})$; namely those whose determinant is real-valued.

In 3+1D, we can define the determinant solely using the constructs of geometric algebra[4].

The determinant of $\mathbf{u}$ is

Definition 7 (3+1D geometric representation of determinant)**.**

$$\det \quad : \qquad \mathcal{G}(\mathbb{R}^{3,1}) \longrightarrow \mathbb{R} \tag{23}$$

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

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

Definition 8 (3+1D Clifford conjugate)**.**

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

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

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

3 Result

3.1 Non-relativistic quantum mechanics

In this subsection, which serves as an introductory example, we recover non-relativistic quantum mechanics using the Lagrange multiplier method and a geometric constraint.

As previously mentionned, the Shannon entropy is applied instead of the Boltzmann entropy to achieve the aforementioned goal.

$$S = -\sum_{q \in \mathbb{Q}} \rho(q) \ln \rho(q) \tag{27}$$

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 *complex geometric constraint*, which is invariant for a complex phase is used. It is defined as

$$\operatorname{tr}\begin{bmatrix}0 & -\overline{b}\\ \overline{b} & 0\end{bmatrix} = \sum_{q\in\mathbb{Q}} \rho(q)\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0,\end{bmatrix} \tag{28}$$

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. In our framework, this phase invariance originates from the trace.

The Lagrangian equation, introduced earlier, that maximizes the entropy subject to the complex geometric constraint is

$$\mathcal{L} = -\sum_{q\in\mathbb{Q}} \rho(q)\ln(q) + \alpha\left(1-\sum_{q\in\mathbb{Q}}\rho(q)\right) + \tau\left(\operatorname{tr}\begin{bmatrix}0 & -\overline{b}\\ \overline{b} & 0\end{bmatrix} - \sum_{q\in\mathbb{Q}}\rho(q)\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{29}$$

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

$$\frac{\partial\mathcal{L}}{\partial\rho(q)} = -\ln\rho(q) - 1 - \alpha - \tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{30}$$

$$0 = \ln\rho(q) + 1 + \alpha + \tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{31}$$

$$\implies \ln\rho(q) = -1 - \alpha - \tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{32}$$

$$\implies \rho(q) = \exp(-1-\alpha)\exp\left(-\tau\operatorname{tr}\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right) \tag{33}$$

$$= \frac{1}{Z(\tau)}\det\exp\left(-\tau\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix}\right), \tag{34}$$

where $Z(\tau)$ is obtained as

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

$$\implies \big(\exp(-1-\alpha)\big)^{-1} = \sum_{q \in \mathbb{Q}} \exp\left(-\tau \operatorname{tr} \begin{bmatrix} 0 & -b(q) \\ b(q) & 0 \end{bmatrix}\right) \tag{36}$$

$$Z(\tau) \coloneqq \sum_{q \in \mathbb{Q}} \det \exp\left(-\tau \begin{bmatrix} 0 & -b(q) \\ b(q) & 0. \end{bmatrix}\right) \tag{37}$$

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

Finally, we obtain

$$\begin{aligned} \rho(\tau, q) &= \frac{1}{Z(\tau)} \det \exp\left(-\tau \begin{bmatrix} 0 & -b(q) \\ b(q) & 0 \end{bmatrix}\right) && (38) \\ &\cong |\exp -i\tau b(q)|^2 && \text{Born rule} \quad (39) \end{aligned}$$

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

$$\rho(q) = \frac{1}{Z} \Big|\exp\big(-itH(q)/\hbar\big)\Big|^2 . \tag{40}$$

or even a more familiar form of

$$\rho(q) = \frac{1}{Z} \big|\psi(q)\big|^2 , \text{ where } \psi(q) = \exp\big(-itH(q)/\hbar\big). \tag{41}$$

With this, we can show that all 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 (41), 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

$$Z = \langle \psi | \psi \rangle . \tag{42}$$

After normalization, the physical states become its unit vectors, and the probability of any particular state is given by

$$\rho(q) = \frac{1}{\langle \psi | \psi \rangle} (\psi(q))^\dagger \psi(q). \tag{43}$$

Finally, any self-adjoint matrix, defined as $\langle \mathbf{O}\psi | \phi \rangle = \langle \psi | \mathbf{O}\phi \rangle$, 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 in its full generality:

$$\operatorname{tr} \overline{\mathbf{M}} = \sum_{q \in \mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{M}(q), \tag{44}$$

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

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

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

$$\mathcal{L} = -\sum_{q \in \mathbb{Q}} \rho(q) \ln(q) + \alpha \left(1 - \sum_{q \in \mathbb{Q}} \rho(q) \right) + \tau \left(\operatorname{tr} \overline{\mathbf{M}} - \sum_{q \in \mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{M}(q) \right), \tag{45}$$

where α and τ are the Lagrange multipliers.

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

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \rho(q)} &= -\ln \rho(q) - 1 - \alpha - \tau \operatorname{tr} \mathbf{M}(q) && (46)\\
0 &= \ln \rho(q) + 1 + \alpha + \tau \operatorname{tr} \mathbf{M}(q) && (47)\\
\implies \ln \rho(q) &= -1 - \alpha - \tau \operatorname{tr} \mathbf{M}(q) && (48)\\
\implies \rho(q) &= \exp(-1 - \alpha) \exp\big(-\tau \operatorname{tr} \mathbf{M}(q)\big) && (49)\\
&= \frac{1}{Z(\tau)} \det \exp\big(-\tau \mathbf{M}(q)\big) && (50)
\end{aligned}$$

where $Z(\tau)$ is obtained as

$$\begin{aligned}
1 &= \sum_{q \in \mathbb{Q}} \exp(-1 - \alpha) \exp\big(-\tau \operatorname{tr} \mathbf{M}(q)\big) && (51)\\
\implies \big(\exp(-1 - \alpha)\big)^{-1} &= \sum_{q \in \mathbb{Q}} \exp\big(-\tau \operatorname{tr} \mathbf{M}(q)\big) && (52)\\
Z(\tau) &:= \sum_{q \in \mathbb{Q}} \det \exp\big(-\tau \mathbf{M}(q)\big) && (53)
\end{aligned}$$

The resulting probability measure is

$$\rho(q, \tau) = \frac{1}{Z(\tau)} \det \exp\big(-\tau \mathbf{M}(q)\big), \tag{54}$$

where

$$Z(\tau)=\sum_{q\in\mathbb{Q}}\det\exp\bigl(-\tau\mathbf{M}(q)\bigr). \tag{55}$$

By defining $\psi(q,\tau):=\exp\bigl(-\tau\mathbf{M}(q)\bigr)$, we can express $\rho(q,\tau)=\det\psi(q,\tau)$, where the determinant acts as a "generalized Born rule," connecting, in this case, a general linear amplitude to a real-valued probability.

The sophistication of the general linear amplitude along with the determinant acting as a "generalized Born rule" will provide a platform for us to support both fundamental physics.

Let us remark that a more general case exists; where a Lagrange multiplier is assigned to each independent entry of the matrix $\mathbf{M}(q)$. In the case, the result would be:

$$\rho(q,\tau)=\frac{1}{Z(\tau)}\det\exp\bigl(-\boldsymbol{\tau}\cdot\mathbf{M}(q)\bigr), \tag{56}$$

where $\boldsymbol{\tau}$ is now a $n\times n$ matrix, and where the $\cdot$ operator assigns the first element of $\boldsymbol{\tau}$ to the first element of $\mathbf{M}(q)$, and so on.

4 Analysis

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

We begin by introducing new groups relevant to the derived probability measurement, then we investigate our result in 2D and 4D.

The 2D definition of the algebra constitutes a special case that is reminiscent of the definitions of ordinary quantum mechanics, yet includes gravity. The 3+1D case is significantly more sophisticated than the 2D case and is elucidated immediately after the 2D case analysis.

4.1 Axiomatic definition of the algebra in 2D

Let $\mathbb{V}$ be an m-dimensional vector space over $\mathcal{G}(\mathbb{R}^2)$.

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

$$\begin{aligned}\langle\cdot,\cdot\rangle\quad:\qquad \mathbb{V}\times\mathbb{V}&\longrightarrow\mathcal{G}(\mathbb{R}^2)\\ \langle\mathbf{u},\mathbf{v}\rangle&\longmapsto\mathbf{u}^\ddagger\mathbf{v}\end{aligned} \tag{57}$$

is positive-definite for $\boldsymbol{\psi}$, such that $\langle\boldsymbol{\psi},\boldsymbol{\psi}\rangle>0$

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

$$\rho(\psi(q), \boldsymbol{\psi}) = \frac{1}{\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle} \psi(q)^{\ddagger} \psi(q) \tag{58}$$

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:

$$\sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q), \boldsymbol{\psi}) = 1 \tag{59}$$

- $\boldsymbol{\psi}$ 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.

$$\sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q), \mathbf{T}\boldsymbol{\psi}) = \sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q), \boldsymbol{\psi}) = 1 \tag{60}$$

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

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

$$\langle \mathbf{O}\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{O}\mathbf{v} \rangle \tag{61}$$

is called an observable.

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

$$\langle \mathbf{O} \rangle = \frac{1}{\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle} \langle \mathbf{O}\boldsymbol{\psi}, \boldsymbol{\psi} \rangle \tag{62}$$

4.2 Geometric self-adjoint operator in 2D

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. It is defined as

$$\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle = \langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle \tag{63}$$

$\forall \boldsymbol{\phi} \in \mathbb{V}$ and $\forall \boldsymbol{\psi} \in \mathbb{V}$.

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 $\mathcal{G}(\mathbb{R}^2)$.

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

$$\begin{aligned} 2\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle &= (\mathbf{o}_{00}\boldsymbol{\phi}_1 + \mathbf{o}_{01}\boldsymbol{\phi}_2)^{\ddagger}\boldsymbol{\psi}_1 + \boldsymbol{\psi}_1^{\ddagger}(\mathbf{o}_{00}\boldsymbol{\phi}_1 + \mathbf{o}_{01}\boldsymbol{\phi}_2) \\ &\quad + (\mathbf{o}_{10}\boldsymbol{\phi}_1 + \mathbf{o}_{11}\boldsymbol{\phi}_2)^{\ddagger}\boldsymbol{\psi}_2 + \boldsymbol{\psi}_2^{\ddagger}(\mathbf{o}_{10}\boldsymbol{\phi}_1 + \mathbf{o}_{11}\boldsymbol{\phi}_2) && (64) \\ &= \boldsymbol{\phi_1}^{\ddagger}\mathbf{o}_{00}^{\ddagger}\boldsymbol{\psi}_1 + \boldsymbol{\phi}_2^{\ddagger}\mathbf{o}_{01}^{\ddagger}\boldsymbol{\psi}_1 + \boldsymbol{\psi}_1^{\ddagger}\mathbf{o}_{00}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_1^{\ddagger}\mathbf{o}_{01}\boldsymbol{\phi}_2 \\ &\quad + \boldsymbol{\phi}_1^{\ddagger}\mathbf{o}_{10}^{\ddagger}\boldsymbol{\psi}_2 + \boldsymbol{\phi}_2^{\ddagger}\mathbf{o}_{11}^{\ddagger}\boldsymbol{\psi}_2 + \boldsymbol{\psi}_2^{\ddagger}\mathbf{o}_{10}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^{\ddagger}\mathbf{o}_{11}\boldsymbol{\phi}_2 && (65) \end{aligned}$$

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

$$\begin{aligned} 2\langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle &= \boldsymbol{\phi}_1^{\ddagger}(\mathbf{o}_{00}\boldsymbol{\psi}_1 + \mathbf{o}_{01}\boldsymbol{\psi}_2) + (\mathbf{o}_{00}\boldsymbol{\psi}_1 + \mathbf{o}_{01}\boldsymbol{\psi}_2)^{\ddagger}\boldsymbol{\phi}_1 \\ &\quad + \boldsymbol{\phi}_2^{\ddagger}(\mathbf{o}_{10}\boldsymbol{\psi}_1 + \mathbf{o}_{11}\boldsymbol{\psi}_2) + (\mathbf{o}_{10}\boldsymbol{\psi}_1 + \mathbf{o}_{11}\boldsymbol{\psi}_2)^{\ddagger}\boldsymbol{\phi}_1 && (66) \\ &= \boldsymbol{\phi}_1^{\ddagger}\mathbf{o}_{00}\boldsymbol{\psi}_1 + \boldsymbol{\phi}_1^{\ddagger}\mathbf{o}_{01}\boldsymbol{\psi}_2 + \boldsymbol{\psi}_1^{\ddagger}\mathbf{o}_{00}^{\ddagger}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^{\ddagger}\mathbf{o}_{01}^{\ddagger}\boldsymbol{\phi}_1 \\ &\quad + \boldsymbol{\phi}_2^{\ddagger}\mathbf{o}_{10}\boldsymbol{\psi}_1 + \boldsymbol{\phi}_2^{\ddagger}\mathbf{o}_{11}\boldsymbol{\psi}_2 + \boldsymbol{\psi}_1^{\ddagger}\mathbf{o}_{10}^{\ddagger}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^{\ddagger}\mathbf{o}_{11}^{\ddagger}\boldsymbol{\phi}_1 && (67) \end{aligned}$$

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

$$\begin{aligned} \mathbf{o}_{00}^{\ddagger} &= \mathbf{o}_{00} && (68) \\ \mathbf{o}_{01}^{\ddagger} &= \mathbf{o}_{10} && (69) \\ \mathbf{o}_{10}^{\ddagger} &= \mathbf{o}_{01} && (70) \\ \mathbf{o}_{11}^{\ddagger} &= \mathbf{o}_{11}. && (71) \end{aligned}$$

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

$$\mathbf{O}^{\ddagger} = \mathbf{O}, \tag{72}$$

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

4.3 Geometric spectral theorem in 2D

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

Consider

$$\mathbf{O} = \begin{bmatrix} a_{00} & a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12} \\ a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} & a_{11} \end{bmatrix}, \tag{73}$$

Then $\mathbf{O}^{\ddagger}$ is

$$\mathbf{O}^{\ddagger} = \begin{bmatrix} a_{00} & a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12} \\ a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} & a_{11} \end{bmatrix}, \tag{74}$$

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

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

The eigenvalues are obtained as

$$0 = \det(\mathbf{O} - \lambda\mathbf{I}) = \det \begin{bmatrix} a_{00} - \lambda & a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12} \\ a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} & a_{11} - \lambda \end{bmatrix}, \tag{75}$$

This implies that

$$0 = (a_{00} - \lambda)(a_{11} - \lambda) - (a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12})(a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} + a_{11}) \tag{76}$$

$$0 = (a_{00} - \lambda)(a_{11} - \lambda) - (a^2 - x^2 - y^2 + b^2), \tag{77}$$

Finally,

$$\lambda = \{\frac{1}{2}\left(a_{00} + a_{11} - \sqrt{(a_{00} - a_{11})^2 + 4(a^2 - x^2 - y^2 + b^2)}\right), \tag{78}$$

$$\frac{1}{2}\left(a_{00} + a_{11} + \sqrt{(a_{00} - a_{11})^2 + 4(a^2 - x^2 - y^2 + b^2)}\right)\} \tag{79}$$

Notably, in the case where $a_{00} - a_{11} = 0$, the roots would be complex if $a^2 - x^2 - y^2 + b^2 < 0$. Is this possible? Note that the determinant of real matrices must be greater than zero because of the exponential mapping to the orientation-preserving general linear group:

$$\exp \mathbf{M}(\mathbb{R}^n) \to \mathrm{GL}^+(n, \mathbb{R}) \tag{80}$$

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 an observable in the typical sense of an observable whose eigenvalues are real-valued.

4.4 Left action in 2D

A left action on the wavefunction $\mathbf{T}\left|\psi\right\rangle$ 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

$$\left\langle\psi\right| \mathbf{T}^\ddagger\mathbf{T}\left|\psi\right\rangle = \left\langle\psi|\psi\right\rangle . \tag{81}$$

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

$$\mathbf{T}^\ddagger\mathbf{T} = \mathbf{I}. \tag{82}$$

Let us consider a two-state system.

A general transformation of such a system is represented by

$$\mathbf{T} = \begin{bmatrix} u & v \\ w & x \end{bmatrix}, \tag{83}$$

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

The expression $\mathbf{T}^\ddagger\mathbf{T}$ is

$$\mathbf{T}^\ddagger\mathbf{T} = \begin{bmatrix} v^\ddagger & u^\ddagger \\ w^\ddagger & x^\ddagger \end{bmatrix} \begin{bmatrix} v & w \\ u & x \end{bmatrix} = \begin{bmatrix} v^\ddagger v + u^\ddagger u & v^\ddagger w + u^\ddagger x \\ w^\ddagger v + x^\ddagger u & w^\ddagger w + x^\ddagger x \end{bmatrix} \tag{84}$$

For $\mathbf{T}^\ddagger\mathbf{T} = \mathbf{I}$, the following relations must hold:

$$v^\ddagger v + u^\ddagger u = 1 \tag{85}$$
$$v^\ddagger w + u^\ddagger x = 0 \tag{86}$$
$$w^\ddagger v + x^\ddagger u = 0 \tag{87}$$
$$w^\ddagger w + x^\ddagger x = 1 \tag{88}$$

This is the case if

$$\mathbf{T} = \frac{1}{\sqrt{v^{\ddagger}v + u^{\ddagger}u}} \begin{bmatrix} v & u \\ -e^{\varphi}u^{\ddagger} & e^{\varphi}v^{\ddagger} \end{bmatrix}, \tag{89}$$

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

$$\mathbf{U} = \frac{1}{\sqrt{|a|^2 + |b|^2}} \begin{bmatrix} a & b \\ -e^{i\theta}b^{\dagger} & e^{i\theta}a^{\dagger} \end{bmatrix}. \tag{90}$$

Here $\mathbf{T}$ is the geometric generalization of unitary transformations.

4.5 The Schrödinger equation of 2D geometry

Let us first recall that the standard Schrödinger equation can be derived as follows.

In the bra-ket notation, we recall that a one-parameter group evolves according to the following equation:

$$\exp(-it\mathbf{H}) \left|\psi(0)\right\rangle = \left|\psi(t)\right\rangle. \tag{91}$$

and thus, an infinitesimal displacement of t is:

$$\exp(-i\delta t\mathbf{H}) \left|\psi(\tau)\right\rangle = \left|\psi(\tau + \delta\tau)\right\rangle. \tag{92}$$

Now, we approximate the exponential into a power series as

$$\exp(-i\delta t\mathbf{H}) \left|\psi(\tau)\right\rangle \approx 1 - i\delta t\mathbf{H} \left|\psi(t)\right\rangle. \tag{93}$$

The process is continued as follows

$$(1 - i\delta t\mathbf{H}) \left|\psi(t)\right\rangle = \left|\psi(t + \delta t)\right\rangle \tag{94}$$

$$\left|\psi(\tau)\right\rangle - i\delta t\mathbf{H} \left|\psi(t)\right\rangle = \left|\psi(t + \delta t)\right\rangle \tag{95}$$

$$-i\delta t\mathbf{H} \left|\psi(t)\right\rangle = \left|\psi(t + \delta t)\right\rangle - \left|\psi(t)\right\rangle \tag{96}$$

$$-i\mathbf{H} \left|\psi(t)\right\rangle = \frac{\left|\psi(t + \delta t)\right\rangle - \left|\psi(t)\right\rangle}{\delta t} \tag{97}$$

$$-i\mathbf{H} \left|\psi(t)\right\rangle = \frac{d\left|\psi(t)\right\rangle}{dt}. \tag{98}$$

which is the Schrödinger equation.

Returning to our result, we now eliminated elements of $\mathbf{u} = a + \mathbf{x} + \mathbf{b}$ by posing $a \to 0, \mathbf{x} \to 0$:

$$\mathbf{u} \mid_{a \to 0, \mathbf{x} \to 0} = \mathbf{b} = \mathbf{i} b \tag{99}$$

This reduces $\mathrm{GL}^+(2, \mathbb{R})$ to $\mathrm{SO}(2)$.

With this elimination, the left action matrix $\mathbf{T}$ becomes valued in $\langle \mathcal{G}(\mathbb{R}^2) \rangle_4$, and the Stone theorem on one-parameter groups applies. Consequently, we can write

$$\mathbf{T}(\tau) \mid_{a \to 0, \mathbf{x} \to 0} = \exp(\mathbf{i} \tau \mathbf{O}) \tag{100}$$

where

$$(\mathbf{O}^{\ddagger} = \mathbf{O}) \mid_{a \to 0, \mathbf{x} \to 0} \implies \mathbf{O}^{\dagger} = \mathbf{O} \tag{101}$$

The end result is an equation that is mathematically similar to the Schrödinger equation (98):

$$-\mathbf{i} \mathbf{O} \left| \psi(\tau) \right\rangle = \frac{d \left| \psi(\tau) \right\rangle}{d\tau}, \tag{102}$$

and the wavefunction is $\psi(\tau) = \exp(-\tau \mathbf{i} \mathbf{O})$

The difference with the Schrödinger equation is that here $\mathbf{i}$ is not the imaginary unit, but a rotor in 2D. We recall that $\mathbf{i} = \hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2$ and that rotors $R = \exp\left(\frac{1}{2}\theta \mathbf{i}\right)$ are exponentials of bivectors.

We have thus arrived at a quantum theory of geometry.

This can be visualized as follows:

$$\psi^{\ddagger}(\tau) \hat{\mathbf{x}}_0 \psi(\tau) = \exp(\tau \mathbf{i} \mathbf{B}) \hat{\mathbf{x}}_0 \exp(-\tau \mathbf{i} \mathbf{B}) \tag{103}$$

$$= \exp(\tau \hat{\mathbf{x}}_0 \hat{\mathbf{x}}_1 \mathbf{B}) \hat{\mathbf{x}}_0 \exp(-\tau \hat{\mathbf{x}}_0 \hat{\mathbf{x}}_1 \mathbf{B}) \tag{104}$$

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:

$$\exp(\tau \hat{\mathbf{x}}_0 \hat{\mathbf{x}}_1 \mathbf{B}) \hat{\mathbf{x}}_0 \exp(-\tau \hat{\mathbf{x}}_0 \hat{\mathbf{x}}_1 \mathbf{B}) = \mathbf{e}_0(\tau) \tag{105}$$

Here, we have obtained a geometry-valued Schrödinger equation under an elimination of the elements of $\mathrm{GL}^+(2, \mathbb{R})$ reducing it to $\mathrm{SO}(2)$, and found that it is invariant in the $\mathrm{SO}(2)$ group.

4.5.1 Gravitizing the quantum in 2D

Roger Penrose argued "that the case for gravitizing quantum theory is at least as strong as that for quantizing gravity"[5].

We further stress that a theory which would succeed at gravitizing the quantum does not need to also quantize gravity (and vice-versa). Indeed, it seems reasonable to expect that any consistent singular theory be at most either, but not both.

It is the idea of gravitizing the quantum (rather than quantizing gravity) that is in line with what we are doing here. Indeed, we have made no changes to general relativity. Instead, our entropy maximization procedure produced a wavefunction of the orientation preserving general linear group, whose geometric flexibility exceeds the familiar unitary wavefunction. It is within this extra flexibility that we will find gravity.

In the previous result we have bluntly eliminated elements of the group $\mathrm{GL}^+(2,\mathbb{R})$ reducing it to $\mathrm{SO}(2)$ by posing $a \to 0, \mathbf{x} \to 0$. How important are the terms we have eliminated? What if instead of eliminating them, we perform a structure reduction thus, recovering the SO(2) group as before, but also the space resulting from a quotient bundle?

Let us investigate.

First, we note that in the general case, our wavefunction is valued in curvilinear (arbitrary basis) multivectors $\mathbf{u}$:

$$\mathbf{u} = a + x\mathbf{e}_1 + y\mathbf{e}_2 + b\mathbf{e}_1\mathbf{e}_2 \tag{106}$$

Second, let X^2 be a smooth orientable real-valued manifold in 2D.

We now equip the manifold X^2 with curvilinear $\mathbf{u}$ via the cross product: $X^2 \times \exp\big(\mathcal{G}(\mathbb{R}^2)\big)$. The crossing induces a frame-bundle FX on X^2, whose structure group is in $\mathrm{GL}^+(2,\mathbb{R})$.

X^2 now has the minimal structure (an exponentiated arbitrary-basis multivector is assigned at every point) required for us to define our wavefunction from the entropy-maximization of multivectors at every point on X^2.

The structure group in $\mathrm{GL}^+(2,\mathbb{R})$ of FX can be reduced to SO(2) (yielding the geometric quantum theory of rotations identified in the previous section), then the global section of the quotient bundle FX/SO(2) is a Riemmanian metric on X^2. The connection that preserves the structure SO(2) across the manifold is the Levi-Civita connection.

The frame bundle is a natural bundle that admits general covariant transformations, which are the symmetries of a gravitation theory on X^2.

We stress that the gravitized quantum theory holds before symmetry breaking (in the $\mathrm{GL}^+(2,\mathbb{R})$ group), as well as after symmetry breaking (into gravity + quantum rotations in SO(2)).

We have now exhausted the full geometric expressiveness of theory in 2D.

4.6 Gravitizing the quantum in 2D (another take)

David Hestenes [6] has formulated the wavefunction in the language of geometric algebra in 3+1D.

The 2D version of the geometric algebra formulation of the wavefunction is

$$\psi = \sqrt{\rho}\exp(\mathbf{i}b) \tag{107}$$

such that

$$\psi\psi^{\ddagger} = \sqrt{\rho}\exp(\mathbf{i}b)\sqrt{\rho}\exp(-\mathbf{i}b) = \rho \tag{108}$$

It is obtained from our formalism by eliminating $\mathbf{x}$ from $\mathbf{u}$ by posing $\mathbf{x} \to 0$. Thus, $\mathbf{u}\mid_{\mathbf{x}\to 0}= a + \mathbf{b}$.

The gravitational theory, in this case, would follow from this structure reducition $\mathrm{GL}(2,\mathbb{R})/(\mathbb{R}\times \mathrm{SO}(2))$, yielding the Weyl connection as the connection that preserves this structure, instead of the Levi-Civita connection.

Here, ρ can be seen as the prior (or initial) probability, and the Weyl connection preserves the weight of this prior (in addition to the rotation group) along the manifold.

4.7 Algebra of geometric observables in 3+1D

In this section, the general case in 3+1D is presented.

In 2D, the determinant can be expressed using only the product $\psi^{\ddagger}\psi$, which can be interpreted as an inner product of two vectors. This form allowed us to extend the complex Hilbert space to a *geometric Hilbert space*. We then found that the familiar properties of the complex Hilbert spaces were transferable to the geometric Hilbert space, eventually yielding a two-dimensional gravitized quantum theory in the language of geometric algebra.

Although a similar correspondence exists in 3+1D, it is less recognizable because we need a *quartic-inner-product* (i.e. $\rho = \lfloor\phi^{\ddagger}\phi\rfloor_{3,4}\phi^{\ddagger}\phi$) to produce a real-valued probability in 3+1D.

Thus, in 3+1D, 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* vector space.

Our aim is to find a construction that supports the general linear wavefunction in 3+1D.

To build the right construction, a quartic-inner-product of four terms is devised, which replaces the inner product in the Hilbert space, mapping any four vectors to an element of $\mathcal{G}(\mathbb{R}^{3,1})$, and yielding a complete *quartic-inner-product* vector space.

The familiar quantum mechanical features (linear transformations, observables as matrix or operators, and linear superposition in the probability measure) will be supported in the construction.

Let $\mathbb{V}$ be a m-dimensional vector space over $\mathcal{G}(\mathbb{R}^{3,1})$.

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 quartic-inner-product form

$$\begin{aligned} \langle \cdot, \cdot, \cdot, \cdot \rangle \quad : \qquad & \mathbb{V} \times \mathbb{V} \times \mathbb{V} \times \mathbb{V} \longrightarrow \mathcal{G}(\mathbb{R}^{3,1}) \\ & \langle \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \rangle \longmapsto \sum_{i=1}^{m} \lfloor u_i^\ddagger v_i \rfloor_{3,4} w_i^\ddagger x_i \end{aligned} \tag{109}$$

   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

$$\rho(\psi(q), \boldsymbol{\phi}) = \frac{1}{\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle} \det \phi(q), \tag{110}$$

   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.

$$\sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q), \boldsymbol{\phi}) = 1 \tag{111}$$

- $\boldsymbol{\phi}$ 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):

$$\sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q), \mathbf{T}\boldsymbol{\phi}) = \sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q), \boldsymbol{\phi}) = 1 \tag{112}$$

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

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

$$\langle \mathbf{Ou}, \mathbf{v}, \mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{Ov}, \mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{v}, \mathbf{Ow}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{Ox} \rangle \tag{113}$$

  is called an observable.

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

$$\langle \mathbf{O} \rangle = \frac{\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle}{\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle} \tag{114}$$

4.7.1 Geometric observables in 3+1D

In 4D, an observable must satisfy equation 113. For simplicity, let us take m in equation 109 to be 1. Then explicitly, we have

$$\lfloor (\mathbf{O}u)^{\ddagger} v \rfloor_{3,4} w^{\ddagger} x = \lfloor u^{\ddagger} \mathbf{O} v \rfloor_{3,4} w^{\ddagger} x = \lfloor u^{\ddagger} v \rfloor_{3,4} (\mathbf{O}w)^{\ddagger} x = \lfloor u^{\ddagger} v \rfloor_{3,4} w^{\ddagger} \mathbf{O} x \tag{115}$$

where u_1, v_1, w_1 and x_1 are multivectors.

Let us investigate.

If $\mathbf{O}$ contained a vector, a bivector, a pseudo-vector or a pseudo-scalar, the equality would not satisfy as these terms do not commune with the multivectors of the equality, and thus cannot be factored out. The equality is satisfied iff $\mathbf{O} \in \mathbb{R}$. Indeed, as a real value, $\mathbf{O}$ commutes with all multivectors of the equality, and can be factored out to satisfy the equality.

We thus find in the general 3+1D case that observables are real-valued.

At a first encounter, this may seem restrictive; comparatively, the observables in the 2D case were geometrically-valued $\mathbf{O}^{\ddagger} = \mathbf{O}$. However, as we will see, the geometric expressivity of the observables in 3+1D expands when we reduce the structure.

Let us investigate the consequences of a structure reduction $\mathbf{x} \to 0, \mathbf{v} \to 0$ of the multivectors of equation 115, and analyse $\mathbf{O}$ again. With such reduction, the multivectors become of the form $\mathbf{u} = a + \mathbf{f} + \mathbf{b}$. This increases the potential for commutativity of $\mathbf{O}$. In fact, we find that $\mathbf{O}$ can now contain scalars and pseudo-scalars, as both commutes with all elements $a, \mathbf{f}$ and $\mathbf{b}$. And with this reduction the observable equation reduces to $\mathbf{O}^{\dagger} = \mathbf{O}$, and this obviously contains more geometry than $\mathbf{O} \in \mathbb{R}$.

4.7.2 Invariant transformations in 3+1D

We now identify the invariant transformations of probability measure (which will be useful later).

$$\lfloor (\mathbf{T}u)^{\ddagger}\mathbf{T}v\rfloor_{3,4}(\mathbf{T}w)^{\ddagger}\mathbf{T}x = \lfloor u^{\ddagger}v\rfloor_{3,4}w^{\ddagger}x \tag{116}$$

$$\implies \lfloor u^{\ddagger}\mathbf{T}^{\ddagger}\mathbf{T}v\rfloor_{3,4}w^{\ddagger}\mathbf{T}^{\ddagger}\mathbf{T}x = \lfloor u^{\ddagger}v\rfloor_{3,4}w^{\ddagger}x \tag{117}$$

The measure is invariant when

1. $\mathbf{T}^{\ddagger}\mathbf{T} = \mathbf{I}$
2. $\mathbf{T}^{\ddagger}\mathbf{T} \in \mathbb{C}^{n\times n}$ and $(\mathbf{T}^{\ddagger}\mathbf{T})^{\dagger}\mathbf{T}^{\ddagger}\mathbf{T} = \mathbf{I}$ and $\mathbf{x} \to 0, \mathbf{v} \to 0$.
3. $\mathbf{T} \in \mathbb{C}^{n\times n}$ and $\mathbf{T}^{\dagger}\mathbf{T} = \mathbf{I}$ and $\mathbf{x} \to 0, \mathbf{v} \to 0$.

4.7.3 Geometric observables in 6D

As we have just noted, the observables in 3+1D must satisfy a more constraining equality relation than in 2D, and this reduced the geometric expressivity that such observable could support. Specifically, in 2D the general observable relation was satisfied for $\mathbf{O}^{\ddagger} = \mathbf{O}$ (this captured the full general linear geometry in 2D), and in 3+1D the general case was satisfied only for $\mathbf{O} \in \mathbb{R}$ (with a structure reductions $\mathbf{x} \to 0, \mathbf{v} \to 0$ yielding $\mathbf{O}^{\dagger} = \mathbf{O}$) which is a tiny subset of the geometric potential in 4D.

What happens if we increased the dimensions even further; to 6 and above?

In dimensions 6 or greater, the corresponding observable relation cannot be satisfied at all. To see why, we have to look at the results[7] of Acus, A et al regarding the 6-dimension multivector norm. In their paper, Acus, A et al disclose having done a brute force computer assisted search for a geometric algebra expression for the determinant in 6D dimension, then as conjectured, found no norm defined only via self-products. The norm they found is a linear combination of self-products. The following is a special case of this norm that holds only for a 6D multivector comprised of a scalar and a grade 4 element:

$$s(B) = b_1 B f_5(f_4(B)f_3(f_2(B)f_1(B))) + b_2 B g_5(g_4(B)g_3(g_2(B)g_1(B))) \tag{118}$$

Even in this simplified special case, we can see that attempting to formulate a relation for observables for a linear equation is doomed to fail. Indeed, even the real portion of the observable cannot be extracted out of the equation. We find that for any of the functions f_i and g_i, the coefficient b_1 and b_2 will frustrate the equality:

$$b_1 \mathbf{O} B f_5(f_4(B)f_3(f_2(B)f_1(B))) + b_2 B g_5(g_4(B)g_3(g_2(B)g_1(B))) \tag{119}$$

$$= b_1 B f_5(f_4(B)f_3(f_2(B)f_1(B))) + b_2 \mathbf{O} B g_5(g_4(B)g_3(g_2(B)g_1(B))) \tag{120}$$

Equation 119 and 120 can only be equal if $b_1 = b_2$, however the norm $s(B)$ requires b_1 and b_2 to be different in general. Consequently, the relation for observables in 6D is unsatisfiable even by real numbers.

Thus, in our framework the geometry of 6D leads to the absence of observables. This result is likely to generalize to all dimensions above 6, as the norms involve more sophisticated systems or linear equations as we go higher.

4.7.4 Defective probability measure in 3D and 5D

We can also rule out the 3D and 5D cases.

The probability measure in these dimensions is not real-valued, but complex-valued, and this makes them defective.

In $\mathcal{G}(\mathbb{R}^3)$, the matrix representation of a multivector

$$\mathbf{u} = a + x\sigma_x + y\sigma_y + z\sigma_z + q\sigma_y\sigma_z + v\sigma_x\sigma_z + w\sigma_x\sigma_y + b\sigma_x\sigma_y\sigma_z \tag{121}$$

is

$$\mathbf{u} \cong \begin{bmatrix} a + ib + iw + z & iq - v + x - iy \\ iq + v + x + iy & a + ib - iw - z \end{bmatrix} \tag{122}$$

and the determinant is

$$\det \mathbf{u} = a^2 - b^2 + q^2 + v^2 + w^2 - x^2 - y^2 - z^2 + 2i(ab - qx + vy - wz) \tag{123}$$

The result is a complex-valued probability. Since a probability must be real-valued, the 3D case is defective in our framework and cannot be used.

In $\mathcal{G}(\mathbb{R}^{4,1})$, the algebra is isomorphic to $\mathcal{G}(\mathbb{C}^{3,1})$ and to complex 4×4 matrices. Consequently, the determinant is complex-valued. Hence, the probability is also complex-valued. Consequently, this case is defective and cannot be used in our framework.

4.7.5 Specialness of 4D

Our framework is non-defective only in:

1. 0D: This corresponds to the familiar (and classical) statistical mechanics. The constraints are scalar $\overline{E} = \sum_{q\in\mathbb{Q}} \rho(q)E(q)$, and the probability measure is the Gibbs measure.

2. 1D: This is the non-relativistic quantum mechanical case we recovered in the results section, using the matrix representation of the complex numbers.

3. 2D: This is the geometric quantum theory we have discussed earlier. Gravity exhausted all geometric freedom of the theory, and thus only gravity exists in 2D. There is no leftover geometry for internal gauges.

4. 4D: This is the case we are investigating here. As we will soon see, the gravitization contains leftover geometry which can be used to define a particle physics.

And is defective in:

1. 3D and 5D: the probability measure is complex-valued.

2. 6D and above: no observables satisfy the corresponds observable equation, in the general case.

Based on our model, it should come with little surprise that the geometry of our universe is four-dimensional. 3+1D is simply the largest spacetime which captures all non-defective cases.

4.7.6 Wavefunction

In the David Hestenes' notation[6], the 3+1D wavefunction is expressed as

$$\psi = \sqrt{\rho e^{ib}} R, \tag{124}$$

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.

To recover the David Hestenes' formulation of the wavefunction, it suffises to square our wavefunction and to eliminate the terms $\mathbf{x} \to 0$ and $\mathbf{v} \to 0$:

$$\psi = \phi^2 \mid_{\mathbf{x}\to 0, \mathbf{v}\to 0} = e^{2a+2\mathbf{f}+2\mathbf{b}} = \sqrt{\rho e^{ib}} R \tag{125}$$

Let us analyse this wavefunction.

First, we can see that the terms a and $\mathbf{b}$ commute with $\mathbf{f}$ and with each other. Thus, they can be factored out as

$$e^{2a+2\mathbf{f}+2\mathbf{b}} = e^{2a+2\mathbf{b}} e^{2\mathbf{f}} \tag{126}$$

Second, the term $\mathbf{f}$ can be understood as the exponential map from the $\mathfrak{so}(3,1)$ lie algebra to the Spin$(3,1)$ group; the double covering being realized from the squaring of our wavefunction allowing R and $-R$ to map to the same rotation.

Consequently, the wavefunction represents the exponential map of the foloowing lie algebra

$$\mathbb{R} \oplus \mathfrak{so}(3,1) \oplus \mathfrak{u}(1) \tag{127}$$

which associates to the following group

$$\mathbb{R} \times \mathrm{Spin}(3,1) \times \mathrm{U}(1) = \mathbb{R} \times \mathrm{Spin}^c(3,1) \tag{128}$$

The gravitational theory follows the same way it did in 2D. We cross a world manifold $\mathbf{X}^4$ with an exponentiated arbitrary curvilinear multivector. This induces a frame bundle LX on $\mathbf{X}^4$, allowing us to define our wavefunction on $\mathbf{X}^4$. The structure reduction $\mathrm{GL}^+(4,\mathbb{R})/\mathrm{SO}(3,1)$ followed by a structure lift to $\text{Spin}^c(3,1)$ entails a U(1)-preserving spin connection as the connection that preserves the structure of the wavefunction on the manifold.

If a prior is used, the structure is reduced to $SO(3,1)$ then lifted to $\mathbb{R} \times \text{Spin}^c(3,1)$ and the associated spin connection also preserves $\mathbb{R} \times U(1)$.

Our models holds in unified form in the unbroken general linear symmetry, and after the symmetry is broken down into two parts: a classical gravity and a unitary quantum theory.

4.7.7 Dirac current

David Hestenes defines the Dirac current in the language of geometric algebra as

$$\mathbf{j} = \psi^{\ddagger}\gamma_0\psi \tag{129}$$

This definition holds in our formulation. We now have all the tools to exhaust the remaining geometric freedom of our framework to construct a particle physics.

4.7.8 SU(2) × U(1) group

Our wavefunction transforms as a group $\psi_1\psi_2 = \psi$. The most general transformation of this type that our framework supports is multiplication by the exponentiation of a reduced multivector (i.e. $\mathbf{x} \to 0, \mathbf{v} \to 0$):

$$e^{\mathbf{u}} = \exp(a + \mathbf{f} + \mathbf{b}) \tag{130}$$

We now simply ask, what is the most general multivector $e^{\mathbf{u}}$ which leaves the Dirac current invariant:

$$\psi^{\ddagger}(e^{\mathbf{u}})^{\ddagger}\gamma_0 e^{\mathbf{u}}\psi = \psi^{\ddagger}\gamma_0\psi \iff (e^{\mathbf{u}})^{\ddagger}\gamma_0 e^{\mathbf{u}} = \gamma_0 \tag{131}$$

When is this satisfied?

The bases of the bivector part $\mathbf{f}$ of $\mathbf{u}$ are $\gamma_0\gamma_1, \gamma_0\gamma_2, \gamma_0\gamma_3, \gamma_1\gamma_2, \gamma_1\gamma_3$, 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, and the rest anti-commute; therefore, the rest must be made equal to 0 lest they won't cancel. Finally, the base $\gamma_0\gamma_1\gamma_2\gamma_3$ anti-commutes with γ_0 thus cancels out.

Consequently, the most general exponential multivector of the form $e^{\mathbf{u}}$ where $\mathbf{u} = \mathbf{f} + \mathbf{b}$ which preserves the Dirac current is

$$e^{\mathbf{u}} = \exp(F_{12}\gamma_1\gamma_2 + F_{13}\gamma_1\gamma_3 + F_{23}\gamma_2\gamma_3 + \mathbf{b}) \tag{132}$$

We can rewrite the bivector basis with the pauli matrices

$$\gamma_2\gamma_3 = \mathbf{i}\sigma_x \tag{133}$$
$$\gamma_1\gamma_3 = \mathbf{i}\sigma_y \tag{134}$$
$$\gamma_1\gamma_2 = \mathbf{i}\sigma_z \tag{135}$$
$$\mathbf{b} = \mathbf{i}b \tag{136}$$

With the replacements, we obtain

$$e^{\mathbf{u}} = \exp \mathbf{i}(F_{12}\sigma_z + F_{13}\sigma_y + F_{23}\sigma_x + b) \tag{137}$$

The terms $F_{23}\sigma_x + F_{13}\sigma_y + F_{12}\sigma_z$ and b are responsible for the $SU(2)$ and $U(1)$ symmetries, respectively. The details of this identification process is available in [8, 9]. David Hestenes and later Lasenby constructed the electroweak sector using the geometric algebra associated with such invariance conditions.

4.7.9 SU(3) group

The invariance identified by equation 117 is $\mathbf{T}^{\ddagger}\mathbf{T} = \mathbf{I}$. The identified evolution was bivectorial rather than unitary.

As we did for the SU(2) $\times$ U(1) case, we will ask, in this case, what is the most general bivectorial evolution which leaves the Dirac current invariant:

$$\mathbf{f}^{\ddagger}\gamma_0\mathbf{f} = \gamma_0 \tag{138}$$

where $\mathbf{f}$ is a bivector:

$$\mathbf{f} = F_{01}\gamma_0\gamma_1 + F_{02}\gamma_0\gamma_2 + F_{03}\gamma_0\gamma_3 + F_{23}\gamma_2\gamma_3 + F_{13}\gamma_1\gamma_3 + F_{12}\gamma_1\gamma_2 \tag{139}$$

Explicitly, the expression $\mathbf{f}^{\ddagger}\gamma_0\mathbf{f}$ is

$$\mathbf{f}^{\ddagger}\gamma_0\mathbf{f} = -\mathbf{f}\gamma_0\mathbf{f} = (F_{01}^2 + F_{02}^2 + F_{03}^2 + F_{13}^2 + F_{23}^2 + F_{12}^2)\gamma_0 \tag{140}$$
$$+ (-2F_{02}F_{12} + 2F_{03}F_{13})\gamma_1 \tag{141}$$
$$+ (-2F_{01}F_{12} + 2F_{03}F_{23})\gamma_2 \tag{142}$$
$$+ (-2F_{01}F_{13} + 2F_{02}F_{23})\gamma_3 \tag{143}$$

For the Dirac current to remain invariant, the cross-product must vanish

$$-2F_{02}F_{12} + 2F_{03}F_{13} = 0 \tag{144}$$
$$-2F_{01}F_{12} + 2F_{03}F_{23} = 0 \tag{145}$$
$$-2F_{01}F_{13} + 2F_{02}F_{23} = 0 \tag{146}$$

leaving only

$$\mathbf{f}^{\ddagger}\gamma_0\mathbf{f} = (F_{01}^2 + F_{02}^2 + F_{03}^2 + F_{13}^2 + F_{23}^2 + F_{12}^2)\gamma_0, \tag{147}$$

Finally, $F_{01}^2 + F_{02}^2 + F_{03}^2 + F_{13}^2 + F_{23}^2 + F_{12}^2$ must equal 1.
We note that we can re-write $\mathbf{f}$ as a 3-vector with complex components:

$$\mathbf{f} = (F_{01} + \mathbf{i}F_{23})\gamma_0\gamma_1 + (F_{02} + \mathbf{i}F_{13})\gamma_0\gamma_2 + (F_{03} + \mathbf{i}F_{12})\gamma_0\gamma_3 \tag{148}$$

Then, with the nullification of the cross-product, and the equality of $F_{01}^2 + F_{02}^2 + F_{03}^2 + F_{13}^2 + F_{23}^2 + F_{12}^2$ to unity, we can understand the bivectorial evolution when constrained by the Dirac current to be a realization of the SU(3) group.

The other invariance of equation 117 is unitary invariance, which was already supported in equation 128.

We have now consumed the geometric expressivity of our framework in 3+1D, to produce the SU(2)×U(1) gauge, the SU(3) gauge and the gravitational theory, leaving no room for anything else.

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, and torsion.

It is plausible that an Aharonov–Bohm effect experiment on gravity[10] 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:

$$\det(\mathbf{u} + \mathbf{v}) = \det\mathbf{u} + \det\mathbf{v} + \text{extra-terms} \tag{149}$$

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.

$$\cdot \quad : \quad \mathcal{G}(\mathbb{R}^{m,n}) \times \mathcal{G}(\mathbb{R}^{m,n}) \longrightarrow \mathbb{R} \tag{150}$$

$$\mathbf{u}\cdot\mathbf{v} \longmapsto \frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det\mathbf{u} - \det\mathbf{v}) \tag{151}$$

For example, in 2D, we have

$$\mathbf{u} = a_1 + x_1\mathbf{e}_1 + y_1\mathbf{e}_2 + b_1\mathbf{e}_{12} \tag{152}$$
$$\mathbf{v} = a_2 + x_2\mathbf{e}_1 + y_2\mathbf{e}_2 + b_2\mathbf{e}_{12} \tag{153}$$
$$\implies \mathbf{u}\cdot\mathbf{v} = a_1a_2 + b_1b_2 - x_1x_2 - y_1y_2. \tag{154}$$

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

$$\frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det \mathbf{u} - \det \mathbf{v}) = \frac{1}{2}\Big((\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v}) - \mathbf{u}^\ddagger\mathbf{u} - \mathbf{v}^\ddagger\mathbf{v}\Big) \tag{155}$$
$$= \mathbf{u}^\ddagger\mathbf{u} + \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} + \mathbf{v}^\ddagger\mathbf{v} - \mathbf{u}^\ddagger\mathbf{u} - \mathbf{v}^\ddagger\mathbf{v} \tag{156}$$
$$= \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} \tag{157}$$

- In 3+1D, it is substantially more complex:

$$\frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det \mathbf{u} - \det \mathbf{v}) \tag{158}$$
$$= \frac{1}{2}\Big(\lfloor(\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v})\rfloor_{3,4}(\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v}) - \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} - \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\Big) \tag{159}$$
$$= \frac{1}{2}\Big(\lfloor\mathbf{u}^\ddagger\mathbf{u} + \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} + \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}(\mathbf{u}^\ddagger\mathbf{u} + \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} + \mathbf{v}^\ddagger\mathbf{v}) - \dots\Big) \tag{160}$$

$$\begin{aligned} = {} & \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\ & + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\ & + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\ & + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} - \dots \end{aligned} \tag{161}$$

$$\begin{aligned} = {} & \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\ & + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\ & + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\ & + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} \end{aligned} \tag{162}$$

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.

$$|\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1||\psi_2| \cos(\phi_1 - \phi_2) \tag{163}$$

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 $[\mathbf{f}, \mathbf{b}] = 0$)).

$$\psi = \psi_1 + \psi_2 = e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} \tag{164}$$

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

$$\lfloor \psi^\ddagger \psi \rfloor_{3,4} \psi^\ddagger \psi. \tag{165}$$

Starting with the sub-product

$$\begin{aligned}
\psi^\ddagger \psi &= (e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2})(e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2}) && (166)\\
&= e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_1} e^{-\mathbf{f}_1} e^{\mathbf{b}_1} e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} \\
&\quad + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2} e^{a_1} e^{\mathbf{f}_1} e^{\mathbf{b}_1} + e^{a_2} e^{-\mathbf{f}_2} e^{\mathbf{b}_2} e^{a_2} e^{\mathbf{f}_2} e^{\mathbf{b}_2} && (167)\\
&= e^{2a_1} e^{2\mathbf{b}_1} + e^{2a_2} e^{2\mathbf{b}_2} + e^{a_1 + a_2} e^{\mathbf{b}_1 + \mathbf{b}_2} (e^{-\mathbf{f}_1} e^{\mathbf{f}_2} + e^{-\mathbf{f}_2} e^{\mathbf{f}_1}) && (168)
\end{aligned}$$

The full product is expressed as

$$
\begin{aligned}
\lfloor\psi^\ddagger\psi\rfloor_{3,4}\psi^\ddagger\psi = &\Big(e^{2a_1}e^{-2\mathbf{b}_1} + e^{2a_2}e^{-2\mathbf{b}_2} + e^{a_1+a_2}e^{-\mathbf{b}_1-\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})\Big)\\
&\times\Big(e^{2a_1}e^{2\mathbf{b}_1} + e^{2a_2}e^{2\mathbf{b}_2} + e^{a_1+a_2}e^{\mathbf{b}_1+\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1}\Big)
\end{aligned}
\tag{169}
$$

$$
\begin{aligned}
= &e^{2a_1}e^{-2\mathbf{b}_1}e^{2a_1}e^{2\mathbf{b}_1} + e^{2a_1}e^{-2\mathbf{b}_1}e^{2a_2}e^{2\mathbf{b}_2} + e^{2a_1}e^{-2\mathbf{b}_1}e^{a_1+a_2}e^{\mathbf{b}_1+\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})\\
&+ e^{2a_2}e^{-2\mathbf{b}_2}e^{2a_1}e^{2\mathbf{b}_1} + e^{2a_2}e^{-2\mathbf{b}_2}e^{2a_2}e^{2\mathbf{b}_2} + e^{2a_2}e^{-2\mathbf{b}_2}e^{a_1+a_2}e^{\mathbf{b}_1+\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})\\
&+ e^{a_1+a_2}e^{-\mathbf{b}_1-\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})e^{2a_1}e^{2\mathbf{b}_1}\\
&\quad + e^{a_1+a_2}e^{-\mathbf{b}_1-\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})e^{2a_2}e^{2\mathbf{b}_2}\\
&\quad + e^{a_1+a_2}e^{-\mathbf{b}_1-\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})e^{a_1+a_2}e^{\mathbf{b}_1+\mathbf{b}_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})
\end{aligned}
\tag{170}
$$

$$
= e^{4a_1} + e^{4a_2} + 2e^{2a_1+2a_2}\cos(2b_1 - 2b_2) \tag{171}
$$

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

$$
e^{2a_1}(e^{-\mathbf{b}_1+\mathbf{b}_2} + e^{\mathbf{b}_1-\mathbf{b}_2}) \tag{173}
$$

$$
+ e^{2a_2}(e^{\mathbf{b}_1-\mathbf{b}_2} + e^{-\mathbf{b}_1+\mathbf{b}_2})) \tag{174}
$$

$$
+ e^{2a_1+2a_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})^2 \tag{175}
$$

$$
\begin{aligned}
= &\underbrace{e^{4a_1} + e^{4a_2}}_{\text{sum}} + \underbrace{2e^{2a_1+2a_2}\cos(2b_1 - 2b_2)}_{\text{complex interference}}\\
&+ \underbrace{2e^{a_1+a_2}(e^{2a_1} + e^{2a_2})(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})(\cos(B_1 - B_2)) + e^{2A_1+2A_2}(e^{-\mathbf{f}_1}e^{\mathbf{f}_2} + e^{-\mathbf{f}_2}e^{\mathbf{f}_1})^2}_{\text{deep spinor interference}}
\end{aligned}
\tag{176}
$$

6 Discussion

We have recovered the foundations of quantum mechanics using the tools of statistical mechanics to maximize the entropy under the effect of a geometric constraint. We have replaced the Boltzmann entropy with the Shannon entropy, and this has an impact on the resulting interpretation, which we will now discuss.

In contrast to the multiple interpretations of quantum mechanics, the interpretation of statistical mechanics is singular, free of paradoxes and clearly 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 of 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 a level similar to the laws of physics themselves.

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 of a message of measurements to an expectation value; or more mathematically, an instrument is described by an equality which constrains the entropy to a given exception value.*

Nature allows geometrically richer measurements and instrumentations, which are not possible to express with simple "scalar" or "phase-less" instruments. For instance, a protractor, a boost meter or shear meter 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 multiplied by 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.

As we have seen, the constraint

$$\operatorname{tr}\begin{bmatrix}0 & -\overline{b}\\ \overline{b} & 0\end{bmatrix} = \sum_{q\in\mathbb{Q}} \operatorname{tr}\rho(q)\begin{bmatrix}0 & -b(q)\\ b(q) & 0\end{bmatrix} \tag{177}$$

induces a complex phase invariance into the probability measure $\rho(q) = \left|\exp\bigl(-i\tau b(q)\bigr)\right|^2$, which gives rise to the Born rule and wavefunction.

Moreover, the constraint

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

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 spacetime events.

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[11]) of measurements.

The receipt of such a message by say, an observer, carries information; it is interpreted as the registration of a "click"[12] 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

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

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

$$\psi \implies \frac{P_n |\psi\rangle}{\sqrt{\langle\psi| P_n |\psi\rangle}} \tag{180}$$

The measurement-collapse problem is, in our framework, 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 is mathematically consistent with this understanding.

Finally, as all knowledge of nature comes from the instruments that can be constructed, postulating these instruments (rather than the wavefunction) to be the axioms of physics and 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 related to measurements.

Context 1 (Observability)**.** *Let q be the elements of a statistical ensemble $\mathbb{Q}$. Then $m : \mathbb{Q} \to \mathbb{R}$ is an observable of $\mathbb{Q}$.*

Context 2 (Comprehensibility)**.** *The experience of the observer in nature is defined as the receipt of a message $\mathbf{m} \in (m(\mathbb{Q}))^n$ of n measurements performed on n identical copies of $\mathbb{Q}$.*

Context 3 (Representativeness)**.** *Observations are representative of 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).*

Context 4 (Comprehensiveness)**.** *Observations are comprehensive in the limit: when $|\mathbf{m}| \to \infty$, then $\mathbb{Q}$ is well-defined (i.e., all the elements in $\mathbb{Q}$ are identified).*

Axiom 1 (Geometricity)**.** *A geometric measuring device constrains the entropy of a message of measurement according to the following equality:*

$$\operatorname{tr} \overline{\mathbf{M}} = \sum_{q \in \mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{M}(q) \tag{181}$$

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

Conjecture 1 (Geometric Totality)**.** *The geometric constraint is sufficiently restrictive to represent only the measurements that are possible in nature, yet sufficiently descriptive to represent all such measurements.*

Theorem 1 (The Laws of Physics as a Theorem)**.** *Maximizing the entropy of a message of measurements constrained by a geometric measuring device, yields the model of physics that maximizes the information acquired from said measurements:*

$$\mathcal{L} = -\sum_{q \in \mathbb{Q}} \rho(q) \ln \rho(q) + \lambda \left(1 - \sum_{q \in \mathbb{Q}} \rho(q) \right) + \tau \left(\operatorname{tr} \overline{\mathbf{M}} - \sum_{q \in \mathbb{Q}} \rho(q) \operatorname{tr} \mathbf{M}(q) \right). \tag{182}$$

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

$$\rho(q,\tau) = \frac{1}{Z(\tau)} \det \exp\big(-\tau \mathbf{M}(q)\big), \tag{183}$$

where

$$Z(\tau) = \sum_{q \in \mathbb{Q}} \det \exp\big(-\tau \mathbf{M}(q)\big). \tag{184}$$

Here, the Lagrange multiplier τ represents the one-parameter group evolution of, in the general case, the orientation preserving general linear group $\mathrm{GL}^+(n,\mathbb{R})$*.*

Our framework work only in 0, 1, 2 and 4 dimensions, otherwise either the probability measure is complex-valued or geometric observables fail to exist. Gravity is manifest from the $\mathrm{GL}^+(n,\mathbb{R})$ *group undergoing symmetry breaking to the Lorentz group* $\mathrm{SO}(n-1,1)$ *in 2D or* $\mathrm{Spim}^c(3,1)$ *group in 4D, yielding a gravitational theory from the Levi-Civita or Spin connection. Furthermore, in the 4D case the invariant evolution must satisfy the Dirac current, yielding the local gauges of the standard model.*

The setup is able to accommodate as a quantum theory the full unbroken $\mathrm{GL}^+(n,\mathbb{R})$ *group, as well as its symmetry breaking into the gravitational field and the local gauges of the standard model.*

7 Conclusion

In this paper, we proposed a geometric constraint, which is used to maximize the Shannon entropy. This geometric constraint allows us to derive a probability measure that supports a geometry richer than what was previously used in statistical physics. 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 is derived, and the Born rule is extended to the determinant. The framework produces a non-defective model for the 0D, 1D, 2D and 4D. 4D stands out as the space which is large enough to include all non-defect variations. A gravitational theory results from the $\mathrm{GL}^+(4,\mathbb{R})$ group undergoing symmetry breaking to the $\mathrm{Spin}^c(3,1)$ group. Breaking the symmetry of the general linear wavefunction into the $\mathrm{Spin}^c(3,1)$ group reduces the quantum theory to the $\mathrm{SU}(2)\times$ $\mathrm{U}(1)$ and $\mathrm{SU}(3)$ for its invariant transformations. Finally, an interpretation of quantum mechanics, i.e., 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 (general linear) wavefunction accompanied by the (general linear) Born rule.

8 Statements and Declarations

The author declares no competing interests. The authors did not receive support from any organization for the submitted work.

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. Geometric (clifford) algebra and its applications. *arXiv preprint math/0605280*, 2006.

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

[5] Roger Penrose. On the gravitization of quantum mechanics 1: Quantum state reduction. *Foundations of Physics*, 44:557–575, 2014.

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

[7] A Acus and A Dargys. Inverse of multivector: Beyond p+ q= 5 threshold. *arXiv preprint arXiv:1712.05204*, 2017.

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

[9] 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.

[10] 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.

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

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