An Optimization Problem Whose Solutions Are the Theories of Everything

Alexandre Harvey-Tremblay[†]

[†]Independent scientist, aht@protonmail.ch
April 19, 2023

Abstract

In modern theoretical physics, the laws of physics are formulated as axioms (e.g., the Dirac–Von Neumann axioms, the Wightman axioms, and Newton's laws of motion). While axioms in modern logic hold true merely by definition, the laws of physics are entailed by measurements. This entailment creates an opportunity to *derive* (rather than to postulate) the laws of physics. We propose a maximization problem on the quantity of information associated with the construction by the observer of a message of realized measurements from an ensemble of possible measurements; that is, physics is the solution that provably makes realized measurements maximally informative to the observer. For linear measurements, solving the optimization problem mechanically produces the optimal Hilbert space that accommodates them. Specifically for general linear measurements in 4D, the solution is a Hilbert space that includes gravity for fermions and bosons from the quotient bundle FX/Spin^c(3,1), electromagnetism from the U(1)-bundle, and the standard model from the gauge group SU(3)xSU(2)xU(1). The solution is a remarkable fit to the universe we inhabit. What about higher dimensions? In general, the solution fails to admit normalizable observables above 4 dimensions, suggesting an intrinsic limit to the dimensionality of observable geometry, and by association of spacetime.

1 Introduction

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

In this methodology, it is the theorems (predictions) of the theory that are used (in concert with experiments) to invalidate its axioms (laws).

In logic, however, axioms define what is true in a theory. It follows obviously that its theorems cannot invalidate them.

We thus identified a first dissimilarity between the use of axioms in logic and their use in physics.

As another even more significant dissimilarity (perhaps even fatal), we note that the laws of physics are entailed by *something* (i.e. measurements) whereas axioms are entailed by *nothing* (i.e. they are held to be true by definition).

Motivated by these dissimilarities, we propose a foundation of physics which posit measurements rather than laws. The structure of measurements then becomes the brute fact (axiom) which entails the laws of physics (theorem), thus eliminating the dissimilarity entirely.

Specifically, the laws of physics will be now derived as the *solution* to an optimization problem.

But what problem, if any, does physics solves optimally?

First, we note that falsification can in general demonstrate scientific fitness (or more precisely, lack thereof), but cannot demonstrate optimality. Indeed, as pure mathematics is unaware of the scientific fitness of any sets of axioms, it apriori cannot assign them a fitness score to rank them. Furthermore, scientific fitness is binary (fit/unfit); it has no provable maximum.

To get a maximum, the first step is to constrain the set of all possible physical theories (an uncountably large space of functions comprising the probability measures) by the structure of all measurements (our axiom); with this constraint, they are now empirical theories and are unable to violate the possible measurements of nature. Then, we score them by their explanatory fitness regarding measurements (which can be ordered), rather than by their scientific fitness (which cannot be ordered). We note that scientific fitness still plays a role, but it is transposed from the laws of physics (the theorem) to the mathematical structure of the measurements (the axiom). Thus, our proposal is a *scientific theory of brute measurements* able to score the empirically-constrained physical theories by their explanatory quality and to prove its own explanatory maximum (which are the laws of physics).

Using a language more conductive for mathematization, we propose to maximize the quantity of information associated with the construction by the observer of a message of realized measurements from an ensemble of possible measurements. In this sense, physics is the solution that provably makes realized measurements maximally informative to the observer.

In a more concrete sense, and in the case of linear measurements, our optimization problem becomes a tool able to mechanically generate the optimal Hilbert space accommodating those measurements. It can even mechanically extend the complex Hilbert space to accommodate all linear measurements, as required. As such it is able to generate the (elusive) Hilbert space structure that supports gravity and the standard model. Specifically, the solution for general linear measurements produces a *geometric* Hilbert space (a superset of complex Hilbert spaces) able to support gravity, and in 4D, it additionally contains the standard model.

Corollaries that follow directly from the solution, such as the mathematical origin of the Born rule, the derivation from first principles of the axioms of quantum physics (thus reducing them to theorems), an identification of the correct interpretation of quantum mechanics (the only interpretation whose mathematical formulation is sufficiently powerful to completely derive the quantum theory it interprets, thus proving its interpretational completeness), and the deprecation of the measurement/collapse problem, are also mechanically included in the solution to the optimization problem.

To define the problem rigorously, we first introduce the key structure that makes our approach possible: the *general linear measurement constraint*. Next, we present its rationale.

The construction of the general linear measurement constraint exploits the connection between geometry and probability via the trace. The trace of a matrix can be understood as the expected eigenvalue multiplied by the vector space dimension, and the eigenvalues as the ratios of the distortion of the linear transformation associated with the matrix[1].

Let $\mathbf{u}$ be a multivector of $\mathcal{G}(\mathbb{R}^{m,n})$ (the geometric algebra of $m+n$ dimensions, defined over the real field) and let $\mathbb{Q}$ be a statistical ensemble. The general linear measurement constraint is:

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

where $d = m+n$, and where $\operatorname{tr}\overline{\mathbf{u}}$ denotes the expectation eigenvalue of the statistically weighted sum of multivectors $\mathbf{u}(q)$, parameterized over ensemble $\mathbb{Q}$.

Since the matrix representation of the multivectors of $\mathcal{G}(\mathbb{R}^2)$ and $\mathcal{G}(\mathbb{R}^{3,1})$ are isomorphic to $\mathbb{M}(2,\mathbb{R})$ and $\mathbb{M}(4,\mathbb{R})$, respectively, we can understand the general linear measurement constraint as a representation of all general linear measurements. The use of multivector merely singles out a preferred geometric representation of said general linear measurements.

We note that the trace of a multivector can be obtained by mapping the multivector to its matrix representation (Section 2), and taking its trace.

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

$$\overline{E} = \sum_{q\in\mathbb{Q}} \rho(q)E(q), \tag{2}$$

which is associated with an energy meter that measures the system's energy and produces a series of energy measurements $E_1, E_2, \dots,$ convergent to an expectation value $\overline{E}$.

Another common constraint is related to the volume:

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

which is associated with a volume meter acting on a system and produces a sequence of measured volumes $V_1, V_2, \dots,$ converging to an expectation value $\overline{V}$.

Moreover, the sum over the statistical ensemble must equal 1, as follows:

$$1 = \sum_{q \in \mathbb{Q}} \rho(q) \tag{4}$$

Using equations (2) and (4), a typical statistical mechanical system is obtained by maximizing the entropy using the corresponding Lagrange equation. The Lagrange multiplier method is expressed as:

$$\mathcal{L}(\rho, \lambda, \beta) = -k_B \sum_{q \in \mathbb{Q}} \rho(q) \ln \rho(q) + \lambda \left(1 - \sum_{q \in \mathbb{Q}} \rho(q)\right) + \beta \left(\overline{E} - \sum_{q \in \mathbb{Q}} \rho(q) E(q)\right), \tag{5}$$

where λ and β are the Lagrange multipliers.

By solving $\frac{\partial \mathcal{L}(\rho,\lambda,\beta)}{\partial \rho} = 0$ for ρ, we obtain the Gibbs measure as:

$$\rho(q, \beta) = \frac{1}{Z(\beta)} \exp\bigl(-\beta E(q)\bigr), \tag{6}$$

where

$$Z(\beta) = \sum_{q \in \mathbb{Q}} \exp\bigl(-\beta E(q)\bigr). \tag{7}$$

In our method, Equation 2, a scalar measurement constraint, is replaced with Equation 1, the general linear measurement constraint. In addition to energy or volume meters, we have protractors, and phase, boost, dilation, spin, and shear meters.

As we found, the general linear measurement constraint is compatible with the full machinery of statistical physics. The probability measure resulting from entropy maximization will preserve the expectation eigenvalue of these transformations up to a phase or symmetry group. For instance, based on our entropy maximization procedure, a statistical system measured exclusively using a protractor will carry a local rotation symmetry in the probability of the measured events.

By limiting the definition of constraints to scalar expressions, we believe that statistical physics has failed to capture all measurements available in nature. The general linear measurement constraint redresses the situation and supports the totality of (linear) measurements that are in principle possible.

Finally, it is the relative Shannon entropy (in base e) that we will maximize and not the Boltzmann entropy. The Shannon entropy does not change the mathematical equation for entropy (minus the Boltzmann constant). However, the interpretation will differ. Rather than describing an ergodic system, the solution will relate to the information associated with the construction by the observer of a message of measurements. The laws of physics, as the solution to this optimization problem, provably makes the elements of this message maximally informative to the observer. Further details on the interpretation are provided in the discussion (Section 5).

1.1 Rigorous formulation of the optimization problem

We propose that the laws of physics are derivable by the following context and axiom related exclusively to measurements and their structure.

Context 1 (Ontology)**.** *The experience of the observer in nature is defined as the construction of a message* $\mathbf{m}$ *of* n *measurements:*

$$\mathbf{m} = \mathrm{Dom}(O)^n \tag{8}$$

1. *where* $O\colon \mathbb{Q} \to \mathbb{R}$ *is an observable of* $\mathbb{Q}$*,*
2. *and where* $\mathbb{Q}$ *is a statistical ensemble.*

Axiom 1 (The Fundamental Structure of Measurements)**.** *The general linear measurement constraint is sufficient to describe the structure of all possible measurements in nature (including scalars and geometric measurements):*

$$\frac{1}{d}\operatorname{tr}\overline{\mathbf{u}} = \sum_{q\in\mathbb{Q}} \rho(q)\frac{1}{d}\operatorname{tr}\mathbf{u}(q) \tag{9}$$

where $\operatorname{tr}\mathbf{u}(q)$ *is an observable (i.e.* $O(q) = \operatorname{tr}\mathbf{u}(q)$*), where* $\operatorname{tr}\overline{\mathbf{u}}$ *is its average, and where* $\mathbf{u}$ *corresponds to a multivector of* $\mathcal{G}(\mathbb{R}^{m,n})$ *such that* $d = m + n$*.*

Theorem 1 (The Fundamental Theorem of Physics)**.** *Physics is the solution to an optimization problem that makes the experience of the observer in nature maximally informative. The optimization problem is defined by the following Lagrange equation:*

$$\mathcal{L}(\rho,\lambda,\boldsymbol{\tau}) = -\sum_{q\in\mathbb{Q}}\rho(q)\ln\frac{\rho(q)}{p(q)} + \lambda\left(1-\sum_{q\in\mathbb{Q}}\rho(q)\right) + \left(\frac{1}{n}\operatorname{tr}\boldsymbol{\tau}\odot\overline{\mathbf{u}} - \sum_{q\in\mathbb{Q}}\rho(q)\frac{1}{n}\operatorname{tr}\boldsymbol{\tau}\odot\mathbf{u}(q)\right) \tag{10}$$

where λ *and* $\boldsymbol{\tau}$ *are Lagrange multipliers.*

This optimization problem maximizes the relative Shannon entropy (this is equivalent to maximizing the information associated to the construction of a message) constrained by the general linear measurement constraint.

The notation $\odot$ *designates the Hadamard product, which is an entrywise product. For instance, consider the multivector* $\mathbf{u} = a + x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + b\hat{\mathbf{x}}\hat{\mathbf{y}}$ *of* $\mathcal{G}(\mathbb{R}^2)$*,* *and consider* $\boldsymbol{\tau} = \tau_a + \tau_x\hat{\mathbf{x}} + \tau_y\hat{\mathbf{y}} + \tau_b\hat{\mathbf{x}}\hat{\mathbf{y}}$*, then* $\boldsymbol{\tau} \odot \mathbf{u} = \tau_a a + \tau_x x\hat{\mathbf{x}} + \tau_y y\hat{\mathbf{y}} + \tau_b b\hat{\mathbf{x}}\hat{\mathbf{y}}$*. The Hadamard product allows us to assign a unique Lagrange multiplier to each element of the basis of the multivector, using a compact notation. In* $\mathcal{G}(\mathbb{R}^2)$*, the constraint expands as follows:*

$$
\begin{aligned}
\left(\frac{1}{n}\operatorname{tr}\boldsymbol{\tau}\odot\overline{\mathbf{u}} - \sum_{q\in\mathbb{Q}}\rho(q)\frac{1}{n}\operatorname{tr}\boldsymbol{\tau}\odot\mathbf{u}(q)\right) = \tau_a & \left(\frac{1}{n}\operatorname{tr}\overline{a} - \sum_{q\in\mathbb{Q}}\rho(q)\frac{1}{n}\operatorname{tr}a(q)\right) \\
&+ \tau_x\left(\frac{1}{n}\operatorname{tr}\overline{x}\hat{\mathbf{x}} - \sum_{q\in\mathbb{Q}}\rho(q)\frac{1}{n}\operatorname{tr}x(q)\hat{\mathbf{x}}\right) \\
&+ \tau_y\left(\frac{1}{n}\operatorname{tr}\overline{y}\hat{\mathbf{y}} - \sum_{q\in\mathbb{Q}}\rho(q)\frac{1}{n}\operatorname{tr}y(q)\hat{\mathbf{y}}\right) \\
&+ \tau_b\left(\frac{1}{n}\operatorname{tr}\overline{b}\hat{\mathbf{x}}\hat{\mathbf{y}} - \sum_{q\in\mathbb{Q}}\rho(q)\frac{1}{n}\operatorname{tr}b(q)\hat{\mathbf{x}}\hat{\mathbf{y}}\right)
\end{aligned}
\tag{11}
$$

where $\tau_a, \tau_x, \tau_y, \tau_b$ *are the Lagrange multipliers.*

Comparatively in $\mathcal{G}(\mathbb{R}^{3,1})$*, the constraint would contain 16 Lagrange multipliers (17 if we also count* λ*) because the multivectors of* $\mathcal{G}(\mathbb{R}^{3,1})$ *have 16 basis elements.*

The manuscript is organized as follows: The Methods section introduces tools using geometric algebra, based on the study by Lundholm et al. [3, 4]. Specifically, we use the notion of a determinant for multivectors and the Clifford conjugate for generalizing the complex conjugate. These tools enable the geometric expression of our results.

The Results section presents two solutions for the Lagrange equation. The first applies to an ensemble $\mathbb{Q}$ which is at most countably infinite, and the second applies to the continuum ($\sum \to \int$) where $\mathbb{Q}$ is uncountable.

In the Analysis section we inspect the solution. The optimization problem is able to mechanically produce the correct Hilbert space for the states of the solution. In 0+1D, a complex Hilbert space is recovered, in which the solution is identical to non-relativistic quantum mechanics. To accommodate the states of all general linear measurements in 2D, a *geometric* Hilbert space is obtained, and in 3+1D a *double-copy* geometric Hilbert space is obtained. The last two structures contain gravity, whilst the last one also contain the standard model. Specfically, we show in the general case that the model is a quantum theory whose principal symmetry is generated by the exponential map of multivectors $\exp\mathcal{G}(\mathbb{R}^{3,1})$. As this map is isomorphic to $\exp\mathbb{M}(4,\mathbb{R})$, it acts (up to isomorphism) on the frame bundle FX of a world manifold. In 3+1D, the symmetry breaks into a quantum theory invariant in the $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$ gauge groups, and from the quotient bundle $\mathrm{FX}/\mathrm{Spin}^c(3,1)$ into a theory of gravity and of electromagnetism for charged fermions. Furthermore, we show that the general solution lacks normalizable observables beyond 4D, naturally limiting the dimensionality of spacetime.

Finally, the Discussion section provides an interpretation of quantum mechanics consistent with its newly revealed origin, namely the *maximally informative interpretation*. Central to this interpretation is the understanding that the wavefunction is not fundamental but derived as the solution to an maximization problem on the quantity of information associated to the construction by the observer of a message of realized measurements. It is the only interpretation whose mathematical formulation is sufficiently powerful to exactly derive the quantum theory from the interpretation, and therefore it can prove its *interpretational completeness*.

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, and 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, the corresponding results are $\text{proj}_1(\mathbb{R}^2) = \mathbb{R}_1$ and $\text{proj}_2(\mathbb{R}^2) = \mathbb{R}_2$, respectively.

  The size of a set $\mathbb{X}$ is $|\mathbb{X}|$.

  The symbol $\cong$ indicates an isomorphism, and $\to$ denotes a 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 $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (such that $\mathbf{e}_\nu \cdot \mathbf{e}_\mu = g_{\mu\nu}$), and $\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_n$ (such that $\hat{\mathbf{x}}_\mu \cdot \hat{\mathbf{x}}_\nu = \eta_{\mu\nu}$) if they are orthonormal.

  A geometric algebra of $m+n$D over 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 vector such as $\langle \mathbf{v} \rangle_0 + \langle \mathbf{v} \rangle_1$ form a para-vector; 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.

  Then, 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.

  The notation $\odot$ designates the Hadamard product, which is an entrywise product. For instance, consider the multivector $\mathbf{u} = a + x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + b\hat{\mathbf{x}}\hat{\mathbf{y}}$ of $\mathcal{G}(\mathbb{R}^2)$, and consider $\boldsymbol{\tau} = \tau_a + \tau_x\hat{\mathbf{x}} + \tau_y\hat{\mathbf{y}} + \tau_b\hat{\mathbf{x}}\hat{\mathbf{y}}$, then $\boldsymbol{\tau} \odot \mathbf{u} = \tau_a a + \tau_x x\hat{\mathbf{x}} + \tau_y y\hat{\mathbf{y}} + \tau_b b\hat{\mathbf{x}}\hat{\mathbf{y}}$.

2.2 Geometric representation in 2D

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

A general multivector of $\mathcal{G}(\mathbb{R}^2)$ is given 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 1 (2D geometric representation)**.**

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

Thus, the trace of $\mathbf{u}$ is a.

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 2 (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 3 (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 4 (2D Clifford transpose)**.** *The Clifford transpose is the geometric analog to the conjugate transpose, interpreted as a transpose followed by an element-by-element application of the complex conjugate. Likewise, 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 Geometric representation in 3+1D

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

A general multivector of $\mathcal{G}(\mathbb{R}^{3,1})$ can be written 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 5 (4D geometric representation)**.**

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

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

Thus, the trace of $\mathbf{u}$ is a.

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

The determinant of $\mathbf{u}$ is

Definition 6 (3+1D geometric representation of determinant)**.**

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

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

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

Definition 7 (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 (the plus sign is reversed 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 Results

The Lagrange equation that defines our optimization problem is:

$$\mathcal{L}(\rho, \lambda, \boldsymbol{\tau}) = -\sum_{q \in \mathbb{Q}} \rho(q) \ln \frac{\rho(q)}{p(q)} + \lambda \left(1 - \sum_{q \in \mathbb{Q}} \rho(q)\right) + \left(\frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \overline{\mathbf{u}} - \sum_{q \in \mathbb{Q}} \rho(q) \frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q)\right), \tag{27}$$

where λ and $\boldsymbol{\tau}$ are the Lagrange multipliers, and where $\mathbf{u}(q)$ is an arbitrary multivector of $d = m + n$ dimensions.

To maximize this equation for ρ, we use the criterion $\frac{\partial \mathcal{L}(\rho, \lambda, \boldsymbol{\tau})}{\partial \rho(q)} = 0$ as follows:

$$\frac{\partial \mathcal{L}(\rho, \lambda, \boldsymbol{\tau})}{\partial \rho(q)} = -\ln \frac{\rho(q)}{p(q)} - 1 - \lambda - \frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q) \tag{28}$$

$$0 = \ln \frac{\rho(q)}{p(q)} + 1 + \lambda + \frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q) \tag{29}$$

$$\implies \ln \frac{\rho(q)}{p(q)} = -1 - \lambda - \frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q) \tag{30}$$

$$\implies \rho(q) = p(q) \exp(-1 - \lambda) \exp\left(-\frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q)\right) \tag{31}$$

$$= \frac{1}{Z(\boldsymbol{\tau})} p(q) \det \exp\left(-\frac{1}{d} \boldsymbol{\tau} \odot \mathbf{u}(q)\right) \tag{32}$$

where $Z(\boldsymbol{\tau})$ is obtained as

$$1 = \sum_{q \in \mathbb{Q}} p(q) \exp(-1 - \lambda) \exp\left(-\frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q)\right) \tag{33}$$

$$\implies \big(\exp(-1 - \lambda)\big)^{-1} = \sum_{q \in \mathbb{Q}} p(q) \exp\left(-\frac{1}{d} \operatorname{tr} \boldsymbol{\tau} \odot \mathbf{u}(q)\right) \tag{34}$$

$$Z(\boldsymbol{\tau}) \coloneqq \sum_{q \in \mathbb{Q}} p(q) \det \exp\left(-\frac{1}{d} \boldsymbol{\tau} \odot \mathbf{u}(q)\right) \tag{35}$$

The resulting probability measure is

$$\rho(q, \boldsymbol{\tau}) = \frac{1}{Z(\boldsymbol{\tau})} p(q, 0) \det \exp\left(-\frac{1}{d} \boldsymbol{\tau} \odot \mathbf{u}(q)\right), \tag{36}$$

where

$$Z(\boldsymbol{\tau}) = \sum_{q \in \mathbb{Q}} p(q, 0) \det \exp\left(-\frac{1}{d} \boldsymbol{\tau} \odot \mathbf{u}(q)\right). \tag{37}$$

Finally, we can pose

$$\rho(q,\boldsymbol{\tau}) = \frac{1}{Z(\boldsymbol{\tau})}\det\psi(q,\boldsymbol{\tau}),\ \text{where}\ \psi(q,\boldsymbol{\tau}) = \exp\left(-\frac{1}{d}\boldsymbol{\tau}\odot\mathbf{u}(q)\right)\psi_0(q,0) \quad (38)$$

and where $p(q,0) = \det\psi_0(q,0)$.

Here, the determinant acts as a *geometric* Born rule, connecting, in this case, a *geometric* amplitude to a real-valued probability.

3.1 Continuum case

In his original paper, Claude Shannon did not derive the differential entropy as a theorem: instead, he posited that the discrete entropy ought to be extended by replacing the sum with the integral:

$$-\sum_{q\in\mathbb{Q}}\rho(q)\ln\rho(q) \to -\int_{\mathbb{R}}\rho(x)\ln\rho(x)\,\mathrm{d}x \quad (39)$$

Unfortunately, it was later discovered that the differential entropy is not always positive, and neither is it invariant under a change of parameters. Specifically, it transforms as follows:

$$-\int_{\mathbb{R}}\rho(x)\ln\rho(x)\,\mathrm{d}x \to -\int_{\mathbb{R}}\tilde{\rho}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\ln\left(\tilde{\rho}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\right)\mathrm{d}x \quad (40)$$

$$= -\int_{\mathbb{R}}\tilde{\rho}(y)\ln\left(\tilde{\rho}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\right)\mathrm{d}y \quad (41)$$

Furthermore, due to an argument by Edwin Thompson Jaynes[5, 6], it is known not to be the correct limiting case of the Shannon entropy. Rather, the limiting case is the relative entropy:

$$S = -\int_{\mathbb{R}}\rho(x)\ln\frac{\rho(x)}{p(x)}\,\mathrm{d}x \quad (42)$$

where $p(x)$ is the initial preparation.

The relative entropy, unlike the differential entropy, is invariant with respect to a change of parameter:

$$-\int_{\mathbb{R}}\rho(x)\ln\frac{\rho(x)}{p(x)}\,\mathrm{d}x \to -\int_{\mathbb{R}}\tilde{\rho}(y(x))\frac{dy}{dx}\ln\frac{\tilde{\rho}(y(x))\frac{dy}{dx}}{\tilde{p}(y(x))\frac{dy}{dx}}\,\mathrm{d}x \quad (43)$$

$$= -\int_{\mathbb{R}}\tilde{\rho}(y)\ln\frac{\tilde{\rho}(y)}{\tilde{p}(y)}\,\mathrm{d}y \quad (44)$$

Let us also show that the normalization constraint is invariant with respect to a change of parameter:

$$\int_{\mathbb{R}} \rho(x)\,\mathrm{d}x \to \int_{\mathbb{R}} \tilde{\rho}(y(x)) \frac{dy}{dx}\,\mathrm{d}x \tag{45}$$

$$= \int_{\mathbb{R}} \tilde{\rho}(y)\,\mathrm{d}y \tag{46}$$

Let us now investigate the differential observable. A differential observable is typically formulated as

$$\overline{O} = \int_{\mathbb{R}} O(x)\rho(x)\,\mathrm{d}x \tag{47}$$

But, this expression is not invariant with respect to a change of parameter:

$$\int_{\mathbb{R}} O(x)\rho(x)\,\mathrm{d}x \to \int_{\mathbb{R}} \tilde{O}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\tilde{\rho}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x \tag{48}$$

$$= \int_{\mathbb{R}} \tilde{O}(y)\tilde{\rho}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}y \tag{49}$$

To correct this, we now introduce the relative (with respect to a reference) observable. For instance, if we stretch space by a factor of 2: $x \to 2x$, then the reference must also be stretched by the same amount for the observable to remain invariant. The consequence is that we observe a ratio:

$$\overline{M/R} = \int_{\mathbb{R}} \frac{M(x)}{R(x)}\rho(x)\,\mathrm{d}x \tag{50}$$

Where R is the reference and the ratio $\overline{O} = \overline{U/R}$ is the observable.
We now show that it is invariant with respect to a change of parameter:

$$\int_{\mathbb{R}} \frac{M(x)}{R(x)}\rho(x)\,\mathrm{d}x \to \int_{\mathbb{R}} \frac{\tilde{M}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}}{\tilde{R}(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}}\rho(y(x))\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x \tag{51}$$

$$= \int_{\mathbb{R}} \frac{\tilde{M}(y)}{\tilde{R}(y)}\rho(y)\,\mathrm{d}y \tag{52}$$

With these definitions, the Lagrange equation becomes:

$$\mathcal{L}(\rho,\lambda,\boldsymbol{\tau}) = -\int_{\mathbb{R}} \rho(x)\ln\frac{\rho(x)}{p(x)}\,\mathrm{d}x + \lambda\left(1 - \int_{\mathbb{R}}\rho(x)\,\mathrm{d}x\right) + \left(\frac{1}{d}\operatorname{tr}\boldsymbol{\tau}\odot\frac{\overline{\mathbf{m}}}{\overline{\mathbf{r}}} - \int_{\mathbb{R}}\frac{1}{d}\operatorname{tr}\boldsymbol{\tau}\odot\frac{\mathbf{m}(x)}{\mathbf{r}(x)}\rho(x)\,\mathrm{d}x\right) \tag{53}$$

Maximizing this equation with respect to ρ gives

$$\rho(x, \boldsymbol{\tau}) = \frac{1}{Z(\boldsymbol{\tau})} p(x, 0) \det \exp\left(-\frac{1}{d} \boldsymbol{\tau} \odot \frac{\mathbf{m}(x)}{\mathbf{r}(x)}\right) \tag{54}$$

where

$$Z(\boldsymbol{\tau}) = \int_{\mathbb{R}} p(q, 0) \det \exp\left(-\frac{1}{d} \boldsymbol{\tau} \odot \frac{\mathbf{m}(x)}{\mathbf{r}(x)}\right) \mathrm{d}x \tag{55}$$

The probability measure is now invariant with respect to a change of parameter:

$$\frac{\int_a^b p(x) \det \exp\left(-\frac{1}{d}\boldsymbol{\tau} \odot \frac{\mathbf{m}(x)}{\mathbf{r}(x)}\right) \mathrm{d}x}{\int_{\mathbb{R}} p(x) \det \exp\left(-\frac{1}{d}\boldsymbol{\tau} \odot \frac{\mathbf{m}(x)}{\mathbf{r}(x)}\right) \mathrm{d}x} \to \frac{\int_a^b \tilde{p}(y(x)) \frac{\mathrm{d}y}{\mathrm{d}x} \det \exp\left(-\frac{1}{d}\boldsymbol{\tau} \odot \frac{\tilde{\mathbf{m}}(y(x)) \frac{\mathrm{d}y}{\mathrm{d}x}}{\tilde{\mathbf{r}}(y(x)) \frac{\mathrm{d}y}{\mathrm{d}x}}\right) \mathrm{d}x}{\int_{\mathbb{R}} \tilde{p}(y(x)) \frac{\mathrm{d}y}{\mathrm{d}x} \det \exp\left(-\frac{1}{d}\boldsymbol{\tau} \odot \frac{\tilde{\mathbf{m}}(y(x)) \frac{\mathrm{d}y}{\mathrm{d}x}}{\tilde{\mathbf{r}}(y(x)) \frac{\mathrm{d}y}{\mathrm{d}x}}\right) \mathrm{d}x} \tag{56}$$

$$= \frac{\int_a^b \tilde{p}(y) \det \exp\left(-\frac{1}{d}\boldsymbol{\tau} \odot \frac{\tilde{\mathbf{m}}(y)}{\tilde{\mathbf{r}}(y)}\right) \mathrm{d}x}{\int_{\mathbb{R}} \tilde{p}(y) \det \exp\left(-\frac{1}{d}\boldsymbol{\tau} \odot \frac{\tilde{\mathbf{m}}(y)}{\tilde{\mathbf{r}}(y)}\right) \mathrm{d}y} \tag{57}$$

We can now generalize this result to a manifold.

Let X^4 be a world manifold. Let us also write $\mathbf{u} = \mathbf{m}/\mathbf{r}$. We can write the probability density as follows:

$$\rho(x, y, z, t, \boldsymbol{\tau}) \mid_a^b = \frac{1}{Z(\tau)} \int_a^b p(x, y, z, t, 0) \det \exp\left(-\frac{1}{4} \boldsymbol{\tau} \odot \mathbf{u}(x, y, z, t)\right) \sqrt{|g|}\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z\, \mathrm{d}t \tag{58}$$

where $|g|$ is the absolute value of the determinant of the matrix representation of the metric tensor on the manifold.

Finally, we can define a wavefunction

$$\phi(x, y, z, t, \boldsymbol{\tau}) = \exp\left(-\frac{1}{4} \boldsymbol{\tau} \odot \mathbf{u}(x, y, z, t)\right) \phi_0(x, y, z, t, 0) \tag{59}$$

where $\det\big(\phi_0(x, y, z, t, 0)\big) = p(x, y, z, t, 0)$.

4 Analysis

We first show that in 0+1D, a complex Hilbert space is obtained, that in 2D a *geometric* Hilbert space is obtained, and finally, that in 3+1D a *double-copy* geometric Hilbert space is obtained. We further show that the last two structures include gravity, whilst the last one additionally includes the standard model. As for the first case, it corresponds to non-relativistic quantum mechanics.

4.1 Phase-invariant measurements in 0+1D

In this subsection, which also serves as an introductory example, we recover non-relativistic quantum mechanics using the Lagrange multiplier method and a linear constraint on the relative Shannon entropy.

We recall that in statistical physics the identification of β with the temperature involves the recovery of the Maxwell equations as the equations of states and under the equality $\beta = 1/(k_B T)$. Similarly, here we will identify known laws of physics in which the Lagrange multiplier plays a known role.

As previously mentioned, the relative Shannon entropy (in base e) is applied instead of the Boltzmann entropy to achieve the aforementioned goal.

$$S = -\sum_{q\in\mathbb{Q}} \rho(q) \ln \frac{\rho(q)}{p(q)} \tag{60}$$

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

A *complex measurement constraint*, invariant for a complex phase, is used to overcome this limitation. It is defined[1] as

$$\operatorname{tr}\begin{bmatrix} 0 & -\overline{E} \\ \overline{E} & 0 \end{bmatrix} = \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr}\begin{bmatrix} 0 & -E(q) \\ E(q) & 0, \end{bmatrix} \tag{61}$$

We recall that $\begin{bmatrix} a(q) & -b(q) \\ b(q) & a(q) \end{bmatrix} \cong a(q) + ib(q)$ is the matrix representation of the complex numbers. In terms of multivectors this constraint corresponds to the matrix representation of the pseudoscalar of $\mathcal{G}(\mathbb{R}^{0,1})$.

Similar to energy or volume meters, linear instruments produce a sequence of measurements that converge to an expectation value but with phase invariance. In our framework, this phase invariance originates from the trace.

The Lagrangian equation that describes this optimization problem is:

$$\mathcal{L}(\rho,\lambda,\tau) = -\sum_{q\in\mathbb{Q}} \rho(q) \ln \frac{\rho(q)}{p(q)} + \lambda\left(1 - \sum_{q\in\mathbb{Q}} \rho(q)\right) + \tau\left(\operatorname{tr}\begin{bmatrix} 0 & -\overline{E} \\ \overline{E} & 0 \end{bmatrix} - \sum_{q\in\mathbb{Q}} \rho(q) \operatorname{tr}\begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix}\right) \tag{62}$$

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

[1] We may wonder why we take $n = 1$ (in Equation 1) if the matrix is 2×2. Here, we only use the imaginary part of the complex numbers $a + ib\,|_{a\to 0} = ib$, making the constraint one-dimensional.

$$\frac{\partial \mathcal{L}(\rho,\lambda,\tau)}{\partial \rho(q)} = -\ln \frac{\rho(q)}{p(q)} - 1 - \lambda - \tau \operatorname{tr} \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \tag{63}$$

$$0 = \ln \frac{\rho(q)}{p(q)} + 1 + \lambda + \tau \operatorname{tr} \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \tag{64}$$

$$\implies \ln \frac{\rho(q)}{p(q)} = -1 - \lambda - \tau \operatorname{tr} \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \tag{65}$$

$$\implies \rho(q) = p(q) \exp(-1-\lambda) \exp \left(-\tau \operatorname{tr} \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \right) \tag{66}$$

$$= \frac{1}{Z(\tau)} p(q) \det \exp \left(-\tau \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \right), \tag{67}$$

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

$$1 = \sum_{q \in \mathbb{Q}} p(q) \exp(-1-\lambda) \exp \left(-\tau \operatorname{tr} \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \right) \tag{68}$$

$$\implies \big(\exp(-1-\lambda)\big)^{-1} = \sum_{q \in \mathbb{Q}} p(q) \exp \left(-\tau \operatorname{tr} \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \right) \tag{69}$$

$$Z(\tau) := \sum_{q \in \mathbb{Q}} p(q) \det \exp \left(-\tau \begin{bmatrix} 0 & -E(q) \\ E(q) & 0. \end{bmatrix} \right) \tag{70}$$

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

$$\rho(q,\tau) = \frac{1}{Z(\tau)} p(q,0) \det \exp \left(-\tau \begin{bmatrix} 0 & -E(q) \\ E(q) & 0 \end{bmatrix} \right) \tag{71}$$

$$\cong p(q,0) |\exp -i\tau E(q)|^2 \tag{72}$$

With the equality $\tau = t/\hbar$ (analogous to $\beta = 1/(k_B T)$) we recover the familiar form of

$$\rho(q,t) = \frac{1}{Z(t)} p(q,0) \Big| \exp\big(-itE(q)/\hbar\big) \Big|^2 . \tag{73}$$

or in general

$$\rho(q,t) = \frac{1}{Z} \big|\psi(q,t)\big|^2, \text{ where } \psi(q,t) = \exp\big(-itE(q)/\hbar\big)\psi_0(q,0). \tag{74}$$

and where $|\psi_0(q,0)|^2 = p(q,0)$ is the initial preparation.

The time t here emerges as a Lagrange multiplier, which is the same manner in which T, the temperature, emerges in ordinary statistical mechanics.

We can show that the Dirac Von–Neumann axioms and the Born rule are satisfied, revealing the possible origin of quantum mechanics as the solution to an optimization problem on the entropy of measurements.

To do so, we identify the wavefunction as a vector of a complex Hilbert space, and the partition function as its inner product, expressed as:

$$Z = \langle \psi | \psi \rangle \,. \tag{75}$$

As the solution is automatically normalized by the entropy-maximization procedure, the physical states are associated with the unit vectors, and the probability of any particular state is given by

$$\rho(q,t) = \frac{1}{\langle \psi | \psi \rangle} (\psi(q,t))^\dagger \psi(q,t). \tag{76}$$

As the solution is invariant under unitary transformations, it can be transformed out of its eigenbasis, and the energy $E(q)$ is in general represented by a Hamiltonian operator as follows:

$$\big|\psi(t)\big\rangle = \exp\big(-it\mathbf{H}/\hbar\big)\,\big|\psi(0)\big\rangle \tag{77}$$

Any self-adjoint operator, 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.

The dynamics are governed by the Schrödinger equation. Indeed, it suffices to take the time derivative

$$\begin{aligned}
\frac{\partial}{\partial t}\big|\psi(t)\big\rangle &= \frac{\partial}{\partial t}(\exp\big(-it\mathbf{H}/\hbar\big)\,\big|\psi(0)\big\rangle) && (78)\\
&= -i\mathbf{H}/\hbar \exp\big(-it\mathbf{H}/\hbar\big)\,\big|\psi(0)\big\rangle && (79)\\
&= -i\mathbf{H}/\hbar\,\big|\psi(t)\big\rangle && (80)\\
\implies \mathbf{H}\,\big|\psi(t)\big\rangle &= i\hbar\frac{\partial}{\partial t}\big|\psi(t)\big\rangle && (81)
\end{aligned}$$

which is the Schrödinger equation.

Finally, the measurement postulate is imported directly from $\rho(q,\boldsymbol{\tau})$ being a probability measure of statistical mechanics like any other; as it is parametrized over $\mathbb{Q}$, it describes the probability of finding the state at parametrization q upon measurement (in the continuum case, this is a Dirac delta).

Consequently, <u>all</u> axioms of non-relativistic quantum mechanics (including the Born rule and measurement postulate) have been reduced to a specific solution of our optimization problem (i.e. it has been reduced to a single axiom).

4.2 Is $\boldsymbol{\tau}$ always time?

We have to be careful associating τ with time, as this will not true in general. The association worked in the previous example, but this was due to a number of coincidences.

Let us investigate.

First, let us note that given the Schrödinger equation, an operator $\mathbf{U}$ is a symmetry if it commutes with the hamiltonian : $[\mathbf{H}, \mathbf{U}] = 0$.

Second, we recall that in 2D, $\boldsymbol{\tau}$ contains four Lagrange multipliers, and in 3+1D, it contains sixteen. Obviously, our solution does not suggest that there are four time dimensions in 2D, and sixteen such dimensions in 3+1D. When discussing events in spacetime, the wavefunction will already be parametrized in spacetime as $\psi(x, y, z, t, \boldsymbol{\tau})$. Since t is already in there, it makes no sense to consider τ to also represent time.

What then is the role of $\boldsymbol{\tau}$ in general?

Each element of $\boldsymbol{\tau}$ corresponds to the one-parameter group that generates a dynamical evolution governed by the Schrödinger equation. In 2D, four Schrödinger equations will be generated, whereas in 3+1D, we will have sixteen Schrödinger equations. They define the dynamics that continuously connect the configurations of spacetime, whilst respecting global symmetries and conservation laws. In neither of these cases will τ represent time. The global symmetries and conservation laws will be induced by the geometric properties of the system.

Let us see in more details.

In 2D, we have 4 Lagrange multipliers $\boldsymbol{\tau} = \tau_a + \tau_x \hat{\mathbf{x}} + \tau_y \hat{\mathbf{y}} + \tau_b \hat{\mathbf{x}}\hat{\mathbf{y}}$. The dynamics of the wavefunction

$$\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle = \exp\left(-\frac{1}{2}\boldsymbol{\tau} \odot \mathbf{u}(x, y)\right)\left|\psi_0(x, y, 0, 0, 0, 0)\right\rangle \tag{82}$$

are given in the form of four partial differential equations:

$$-\frac{1}{2}a\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle = \frac{\partial}{\partial \tau_a}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle \tag{83}$$

$$-\frac{1}{2}x\hat{\mathbf{x}}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle = \frac{\partial}{\partial \tau_x}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle \tag{84}$$

$$-\frac{1}{2}y\hat{\mathbf{y}}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle = \frac{\partial}{\partial \tau_y}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle \tag{85}$$

$$-\frac{1}{2}b\hat{\mathbf{x}}\hat{\mathbf{y}}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle = \frac{\partial}{\partial \tau_b}\left|\psi(x, y, \tau_a, \tau_x, \tau_y, \tau_b)\right\rangle \tag{86}$$

The Schrödinger-like equations define four relations for conserved quantities:

$$[a, \mathbf{u}] = 0 \tag{87}$$
$$[x\hat{\mathbf{x}}, \mathbf{u}] = 0 \tag{88}$$
$$[y\hat{\mathbf{y}}, \mathbf{u}] = 0 \tag{89}$$
$$[b\hat{\mathbf{x}}\hat{\mathbf{y}}, \mathbf{u}] = 0 \tag{90}$$

The first relation define the set of all $\mathbf{u}$ which are conserved under the generator of dilation transformations (since a is a scalar, this case exceptionally comprises all $\mathbf{u}$ in $\mathcal{G}(\mathbb{R}^2)$), the second and third relate to the generators of shear transformations, and the fourth to the generator of rotational transformations.

Why does τ associates to t in non-relativistic quantum mechanics? This is caused by three coincidences. In non-relativistic quantum mechanics, the wavefunction is parametrized in (x, y, z), but not in t (the first coincidence). Furthermore, in this case only one Schrödinger equation is required to connect the different configurations of space (the second coincidence). Finally, the pseudoscalar happens to generate the geometric symmetry of the Hamiltonian (third coincidence). Because of the three coincidences, τ can be associated with time simply by choosing the factor of the pseudoscalar to be the Hamiltonian.

Our general interpretation of the Schrödinger equation as continuously connecting two configurations of spacetime also holds in the non-relativistic case; indeed, the Schrödinger equation connects a configuration of space $\psi_0(x, y, z)$ to another $\psi(x, y, z)$ using the one-parameter group generated by τ.

To further clarify the dynamics let us contrast it to the dynamics commonly found in relativity. For instance, an observer or particle in 3+1D might still follow a path l parametrizing the coordinates (x, y, z, t) as $(x(l), y(l), z(l), t(l))$ and consequently still be the subject of a dynamical equation such as the Dirac equation. However, such dynamics would be defined within a configuration of spacetime, and this is not the same as the dynamics that generates the global symmetries and conservation laws via the four Schrödinger equations of 2D, or the sixteen of 3+1D.

4.3 Geometric Hilbert space in 2D

We now attack the 2D case. We recall that the determinant in 2D can be expressed as $\det \mathbf{u} = \mathbf{u}^\ddagger \mathbf{u}$, where $\mathbf{u}^\ddagger$ is the Clifford conjugate of $\mathbf{u}$. This allows us to use a notation similar to the bra-ket notation used in physics. It also allows us to represent an inner product analogously to the complex norm for complex Hilbert spaces.

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

$$\begin{aligned} \langle\cdot,\cdot\rangle \quad : \qquad \mathbb{V}\times\mathbb{V} &\longrightarrow \mathcal{G}(\mathbb{R}^2) \\ \langle \mathbf{u},\mathbf{v}\rangle &\longmapsto \mathbf{u}^\ddagger\mathbf{v} \end{aligned} \tag{91}$$

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

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

$$\rho(\psi(q)) = \frac{1}{\langle\boldsymbol{\psi},\boldsymbol{\psi}\rangle}\psi(q)^\ddagger\psi(q) \tag{92}$$

is either positive or equal to zero.

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:

$$\sum_{\psi(q)\in\boldsymbol{\psi}} \rho(\psi(q)) = 1 \tag{93}$$

- $\boldsymbol{\psi}$ is called a *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.
- $\rho(q)$ 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.

$$\langle\mathbf{T}\boldsymbol{\psi},\mathbf{T}\boldsymbol{\psi}\rangle = \langle\boldsymbol{\psi},\boldsymbol{\psi}\rangle \tag{94}$$

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

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

$$\langle\mathbf{O}\mathbf{u},\mathbf{v}\rangle = \langle\mathbf{u},\mathbf{O}\mathbf{v}\rangle \tag{95}$$

is called an observable.

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

$$\langle\mathbf{O}\rangle = \frac{1}{\langle\boldsymbol{\psi},\boldsymbol{\psi}\rangle}\langle\mathbf{O}\boldsymbol{\psi},\boldsymbol{\psi}\rangle \tag{96}$$

4.4 Geometric self-adjoint operator in 2D

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

$$\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle = \langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle \tag{97}$$

$\forall \boldsymbol{\phi} \in \mathbb{V}$ 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 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 $\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle$:

$$\begin{aligned} 2\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\psi} \rangle &= (\mathbf{o}_{00}\boldsymbol{\phi}_1 + \mathbf{o}_{01}\boldsymbol{\phi}_2)^\ddagger \boldsymbol{\psi}_1 + \boldsymbol{\psi}_1^\ddagger(\mathbf{o}_{00}\boldsymbol{\phi}_1 + \mathbf{o}_{01}\boldsymbol{\phi}_2) \\ &\quad + (\mathbf{o}_{10}\boldsymbol{\phi}_1 + \mathbf{o}_{11}\boldsymbol{\phi}_2)^\ddagger \boldsymbol{\psi}_2 + \boldsymbol{\psi}_2^\ddagger(\mathbf{o}_{10}\boldsymbol{\phi}_1 + \mathbf{o}_{11}\boldsymbol{\phi}_2) \qquad (98) \\ &= \boldsymbol{\phi_1}^\ddagger \mathbf{o}_{00}^\ddagger \boldsymbol{\psi}_1 + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{01}^\ddagger \boldsymbol{\psi}_1 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{00}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{01}\boldsymbol{\phi}_2 \\ &\quad + \boldsymbol{\phi}_1^\ddagger \mathbf{o}_{10}^\ddagger \boldsymbol{\psi}_2 + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{11}^\ddagger \boldsymbol{\psi}_2 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{10}\boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{11}\boldsymbol{\phi}_2 \qquad (99) \end{aligned}$$

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

$$\begin{aligned} 2\langle \boldsymbol{\phi}, \mathbf{O}\boldsymbol{\psi} \rangle &= \boldsymbol{\phi}_1^\ddagger(\mathbf{o}_{00}\boldsymbol{\psi}_1 + \mathbf{o}_{01}\boldsymbol{\psi}_2) + (\mathbf{o}_{00}\boldsymbol{\psi}_1 + \mathbf{o}_{01}\boldsymbol{\psi}_2)^\ddagger \boldsymbol{\phi}_1 \\ &\quad + \boldsymbol{\phi}_2^\ddagger(\mathbf{o}_{10}\boldsymbol{\psi}_1 + \mathbf{o}_{11}\boldsymbol{\psi}_2) + (\mathbf{o}_{10}\boldsymbol{\psi}_1 + \mathbf{o}_{11}\boldsymbol{\psi}_2)^\ddagger \boldsymbol{\phi}_1 \qquad (100) \\ &= \boldsymbol{\phi}_1^\ddagger \mathbf{o}_{00}\boldsymbol{\psi}_1 + \boldsymbol{\phi}_1^\ddagger \mathbf{o}_{01}\boldsymbol{\psi}_2 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{00}^\ddagger \boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{01}^\ddagger \boldsymbol{\phi}_1 \\ &\quad + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{10}\boldsymbol{\psi}_1 + \boldsymbol{\phi}_2^\ddagger \mathbf{o}_{11}\boldsymbol{\psi}_2 + \boldsymbol{\psi}_1^\ddagger \mathbf{o}_{10}^\ddagger \boldsymbol{\phi}_1 + \boldsymbol{\psi}_2^\ddagger \mathbf{o}_{11}^\ddagger \boldsymbol{\phi}_1 \qquad (101) \end{aligned}$$

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

$$\mathbf{o}_{00}^\ddagger = \mathbf{o}_{00} \tag{102}$$
$$\mathbf{o}_{01}^\ddagger = \mathbf{o}_{10} \tag{103}$$
$$\mathbf{o}_{10}^\ddagger = \mathbf{o}_{01} \tag{104}$$
$$\mathbf{o}_{11}^\ddagger = \mathbf{o}_{11}. \tag{105}$$

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

$$\mathbf{O}^\ddagger = \mathbf{O}, \tag{106}$$

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

4.5 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

$$\mathbf{O} = \begin{bmatrix} a_{00} & a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12} \\ a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} & a_{11} \end{bmatrix}, \tag{107}$$

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

$$\mathbf{O}^{\ddagger} = \begin{bmatrix} a_{00} & a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12} \\ a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} & a_{11} \end{bmatrix}, \tag{108}$$

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:

$$0 = \det(\mathbf{O} - \lambda \mathbf{I}) = \det \begin{bmatrix} a_{00} - \lambda & a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12} \\ a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} & a_{11} - \lambda \end{bmatrix}, \tag{109}$$

This implies that

$$0 = (a_{00} - \lambda)(a_{11} - \lambda) - (a - x\hat{\mathbf{x}}_1 - y\hat{\mathbf{x}}_2 - b\hat{\mathbf{x}}_{12})(a + x\hat{\mathbf{x}}_1 + y\hat{\mathbf{x}}_2 + b\hat{\mathbf{x}}_{12} + a_{11}) \tag{110}$$

$$0 = (a_{00} - \lambda)(a_{11} - \lambda) - (a^2 - x^2 - y^2 + b^2), \tag{111}$$

Finally,

$$\lambda = \{\frac{1}{2}\left(a_{00} + a_{11} - \sqrt{(a_{00} - a_{11})^2 + 4(a^2 - x^2 - y^2 + b^2)}\right), \tag{112}$$

$$\frac{1}{2}\left(a_{00} + a_{11} + \sqrt{(a_{00} - a_{11})^2 + 4(a^2 - x^2 - y^2 + b^2)}\right)\} \tag{113}$$

The roots would be complex if $a^2 - x^2 - y^2 + b^2 < 0$. Since $a^2 - x^2 - y^2 + b^2$ is the determinant of the multivector, the complex case is ruled out for orientation-preserving multivectors. Consequently, it follows $\mathbf{O}^{\ddagger} = \mathbf{O}$ constitutes an observable with real-valued eigenvalues for orientation-preserving multivectors.

4.6 Invariant transformations 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

$$\left\langle\psi\right|\mathbf{T}^{\ddagger}\mathbf{T}\left|\psi\right\rangle = \left\langle\psi|\psi\right\rangle . \tag{114}$$

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

$$\mathbf{T}^{\ddagger}\mathbf{T} = \mathbf{I}. \tag{115}$$

Let us consider a two-state system, with a general transformation represented by

$$\mathbf{T} = \begin{bmatrix} u & v \\ w & x \end{bmatrix}, \tag{116}$$

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

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

$$\mathbf{T}^{\ddagger}\mathbf{T} = \begin{bmatrix} v^{\ddagger} & u^{\ddagger} \\ w^{\ddagger} & x^{\ddagger} \end{bmatrix} \begin{bmatrix} v & w \\ u & x \end{bmatrix} = \begin{bmatrix} v^{\ddagger}v + u^{\ddagger}u & v^{\ddagger}w + u^{\ddagger}x \\ w^{\ddagger}v + x^{\ddagger}u & w^{\ddagger}w + x^{\ddagger}x \end{bmatrix} \tag{117}$$

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

$$v^{\ddagger}v + u^{\ddagger}u = 1 \tag{118}$$

$$v^{\ddagger}w + u^{\ddagger}x = 0 \tag{119}$$

$$w^{\ddagger}v + x^{\ddagger}u = 0 \tag{120}$$

$$w^{\ddagger}w + x^{\ddagger}x = 1 \tag{121}$$

This is the case if

$$\mathbf{T} = \frac{1}{\sqrt{v^{\ddagger}v + u^{\ddagger}u}} \begin{bmatrix} v & u \\ -e^{\varphi}u^{\ddagger} & e^{\varphi}v^{\ddagger} \end{bmatrix}, \tag{122}$$

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

$$\mathbf{U} = \frac{1}{\sqrt{|a|^2 + |b|^2}} \begin{bmatrix} a & b \\ -e^{i\theta}b^{\dagger} & e^{i\theta}a^{\dagger} \end{bmatrix}. \tag{123}$$

Here $\mathbf{T}$ is the geometric generalization (in 2D) of unitary transformations.

4.7 Gravity in FX/SO(2)

We will now investigate the quotient bundle associated with the structure reduction from $\mathrm{GL}^+(2,\mathbb{R})$ to SO(2).

Let X^2 be a smooth orientable real-valued manifold in 2D. We consider its tangent bundle TX and its associated frame bundle FX. Since X^2 is orientable, its structure group is $\mathrm{GL}^+(2,\mathbb{R})$. The action by our wavefunction, valued in $\exp \mathcal{G}(\mathbb{R}^2) \cong \exp \mathbb{M}(2,\mathbb{R})$ generates $\mathrm{GL}^+(2,\mathbb{R})$, and thus acts on FX. We now consider a reduction of the structure group of FX to SO(2).

Let us begin by investigating the cosets of SO(2) in $\mathrm{GL}^+(2,\mathbb{R})$. Let $g_1 \in \mathrm{GL}^+(2,\mathbb{R})$, $g_2 \in \mathrm{GL}^+(2,\mathbb{R})$ and $s \in \mathrm{SO}(2)$. We now identify the relation $g_2 = g_1 s$. We also note $g_2^T = s^T g_1^T$. Finally, we note the product $g_2 g_2^T = g_1 s s^T g_1^T \implies g_2 g_2^T = g_1 g_1^T$. Since $g_1 g_1^T$ and $g_2 g_2^T$ are symmetric positive-definite 2×2 matrices, one verifies a diffeomorphism between $\mathrm{GL}^+(2,\mathbb{R})/\mathrm{SO}(2)$ and the inner products.

The global section of the quotient bundle FX/SO(2) is a tetrad field $h^a_\mu(x)$ and it associates to a Riemannian metric on X^2 via the identity $g_{\mu\nu} = h^a_\mu h^b_\nu \eta_{ab}$. The connection that preserves the structure SO(2) across the manifold are the metric connections[7], and with the additional requirement of no torsion, the connections reduce to the Levi-Civita connection. It has been shown recently[8] that the Goldstone fields associated with the quotient bundle have enough degrees of freedom to create a metric and a covariant derivative. Finally, the frame bundle is a natural bundle that admits general covariant transformations, which are the symmetries of the gravitation theory on X^2[9]. This is the geometric setting for gravity.

In this work, we have merely maximized the entropy of all possible geometric measurements, and we have arrived, without introducing any other assumptions, at a general linear quantum theory holding in the $\mathrm{GL}^+(2,\mathbb{R})$ group, whose symmetry breaks into a theory of gravity (FX/SO(2)) and into a quantum theory of the special orthogonal group (valued in SO(2)).

4.8 Wavefunction in SO(2)

With its structure reduced to SO(2), we thus arrived at a quantum theory of the special orthogonal group, where the wavefunction defines the action on a vector of the tangent space of the manifold, as follows:

$$\psi^\ddagger(\tau_b)\hat{\mathbf{x}}_0\psi(\tau_b) = \exp\left(\tau_b \frac{1}{2}\mathbf{i}B\right)\hat{\mathbf{x}}_0 \exp\left(-\tau_b\frac{1}{2}\mathbf{i}B\right) \tag{124}$$

$$= \exp\left(\tau_b \frac{1}{2}\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1 B\right)\hat{\mathbf{x}}_0 \exp\left(-\tau_b\frac{1}{2}\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1 B\right) \tag{125}$$

The expression $\exp\left(\tau_b \frac{1}{2}\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1 B\right)\hat{\mathbf{x}}_0 \exp\left(-\tau_b\frac{1}{2}\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1 B\right)$ maps $\hat{\mathbf{x}}_0$ to a curvilinear basis $\mathbf{e}_0$ via the application of the rotor and its reverse:

$$\exp\left(\tau_b \frac{1}{2}\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1 B\right)\hat{\mathbf{x}}_0 \exp\left(-\tau_b \frac{1}{2}\hat{\mathbf{x}}_0\hat{\mathbf{x}}_1 B\right) = \mathbf{e}_0(\tau) \tag{126}$$

Consequently, we have obtained a 2D relativistic wavefunction (with Euclidean signature in this case). This is the 2D version of the David Hestenes' geometric algebra formulation of the relativistic wavefunction. In 3+1D case, we will see that the wavefunction has 6 generators for rotations and boosts, and one generator of a complex phase.

4.9 Metric interference in 2D

We now consider a transformation $\mathbf{T}^\ddagger\mathbf{T} = \mathbf{I}$ and a wavefunction $|\psi\rangle = \begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix}$ such that a multivector $\mathbf{u}$ is mapped to a linear combination of two multivectors. Let us consider this transformation:

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{u}+\mathbf{v}\\ \mathbf{u}-\mathbf{v}\end{bmatrix} \tag{127}$$

We can now investigate the probability:

$$\rho(\mathbf{u}+\mathbf{v}) = \frac{1}{Z}\det(\mathbf{u}+\mathbf{v}),\ \text{where } Z = \det(\mathbf{u}+\mathbf{v}) + \det(\mathbf{u}-\mathbf{v}) \tag{128}$$

We proceed as follows:

$$\begin{aligned}
\det(\mathbf{u}+\mathbf{v}) &= (\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v}) && (129)\\
&= (\mathbf{u}^\ddagger+\mathbf{v}^\ddagger)(\mathbf{u}+\mathbf{v}) && (130)\\
&= (\mathbf{u}^\ddagger\mathbf{u}+\mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u}+\mathbf{v}^\ddagger\mathbf{v}) && (131)\\
&= \det\mathbf{u}+\det\mathbf{v}+\mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u} && (132)\\
&= \det\mathbf{u}+\det\mathbf{v}+\mathbf{u}\cdot\mathbf{v} && (133)
\end{aligned}$$

where we have defined the dot product between multivectors as follows:

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^\ddagger\mathbf{v}+\mathbf{v}^\ddagger\mathbf{u} \tag{134}$$

Since $\det\mathbf{u} > 0$ 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 capable of destructive and constructive interference.

In the case $\mathbf{x} \to 0$, the interference pattern reduces to a form identical to the unitary case:

$$\det\Big(\psi_1 e^{-\tau_b \frac{1}{2}\mathbf{b}_1} + \psi_2 e^{-\tau_b \frac{1}{2}\mathbf{b}_2}\Big) = \det\psi_1 + \det\psi_2 + 2\psi_1\psi_2 e^{-\tau_b \frac{1}{2}\mathbf{b}_1 - \tau_b \frac{1}{2}\mathbf{b}_2} \quad (135)$$

$$= |\psi_1|^2 + |\psi_2|^2 + 2\psi_1\psi_2 e^{-\tau_b \frac{1}{2}\mathbf{b}_1 - \tau_b \frac{1}{2}\mathbf{b}_2} \quad (136)$$

whereas, in the general linear case, we would have

$$\det\left(\psi_1 e^{-\frac{1}{2}\boldsymbol{\tau}\odot(a_1+\mathbf{x}_1+\mathbf{b}_1)} + \psi_2 e^{-\frac{1}{2}\boldsymbol{\tau}\odot(a_2+\mathbf{x}_2+\mathbf{b}_2)}\right) \quad (137)$$

$$= \det\psi_1 + \det\psi_2 + 2\psi_1\psi_2\left(e^{-\frac{1}{2}\boldsymbol{\tau}\odot(a_1+\mathbf{x}_1+\mathbf{b}_1)} + e^{-\frac{1}{2}\boldsymbol{\tau}\odot(a_2+\mathbf{x}_2+\mathbf{b}_2)}\right) \quad (138)$$

which includes non-commutative effects in the interference pattern.

4.10 A *double-copy* geometric Hilbert space in 4D

In 2D, the determinant can be expressed using only the product $\psi^\ddagger\psi$, which can be interpreted as the inner product of two multivectors. 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 2D gravitized quantum theory in the language of geometric algebra.

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

Thus, in 4D, we cannot produce an inner product as in the 2D case. The absence of a satisfactory inner product indicates no Hilbert space in the usual sense of a complete *inner product* vector space.

We aim to find a construction that supports the geometric wavefunction in 4D.

To build the right construction, a double-copy inner product of four terms is devised, superseding 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 *double-copy* inner product vector space — or simply, a *double-copy* Hilbert space.

We note that the construction will be more familiar than it may first appear. Indeed, the familiar quantum mechanical features (linear transformations, unit vectors, and linear superposition in the probability measure, etc.) will be supported in the construction, and just as it did in 2D, it will also here break into a familiar inner-product Hilbert space whose Dirac current is invariant for SU(3)×SU(2)×U(1) and into a theory of gravity and of electromagnetism for charged fermions FX/Spin$^c(3,1)$.

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

A subset of vectors in $\mathbb{V}$ forms a double-copy algebra of observables $\mathcal{A}(\mathbb{V})$ if the following holds:

1. $\forall\boldsymbol{\phi}\in\mathcal{A}(\mathbb{V})$, the double-copy inner product form

$$
\begin{aligned}
\langle \cdot, \cdot, \cdot, \cdot \rangle \quad : \qquad & \mathbb{V} \times \mathbb{V} \times \mathbb{V} \times \mathbb{V} \longrightarrow \mathcal{G}(\mathbb{R}^{3,1}) \\
& \langle \mathbf{u}, \mathbf{w}, \mathbf{y}, \mathbf{z} \rangle \longmapsto \sum_{i=1}^{m} \lfloor u_i^\ddagger w_i \rfloor_{3,4} y_i^\ddagger z_i
\end{aligned} \tag{139}
$$

is positive-definite when $\boldsymbol{\phi} \neq 0$; 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 $\phi(q) \in \boldsymbol{\phi}$, the function

$$
\rho(\phi(q)) = \frac{1}{\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle} \det \phi(q), \tag{140}
$$

is either positive or equal to zero.

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.

$$
\sum_{\phi(q) \in \boldsymbol{\phi}} \rho(\phi(q)) = 1 \tag{141}
$$

- $\boldsymbol{\phi}$ is called a *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.
- $\rho(q)$ 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):

$$
\langle \mathbf{T}\boldsymbol{\phi}, \mathbf{T}\boldsymbol{\phi}, \mathbf{T}\boldsymbol{\phi}, \mathbf{T}\boldsymbol{\phi} \rangle = \langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle \tag{142}
$$

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

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

$$
\langle \mathbf{O}\mathbf{u}, \mathbf{w}, \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{u}, \mathbf{O}\mathbf{w}, \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{u}, \mathbf{w}, \mathbf{O}\mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{u}, \mathbf{w}, \mathbf{y}, \mathbf{O}\mathbf{z} \rangle \tag{143}
$$

is called an observable.

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

$$
\langle \mathbf{O} \rangle = \frac{\langle \mathbf{O}\boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle}{\langle \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi}, \boldsymbol{\phi} \rangle} \tag{144}
$$

4.11 Wavefunction in 3+1D

In the David Hestenes' notation[10], the 3+1D wavefunction is expressed as:

$$\psi = \sqrt{\rho e^{-ib}}R, \tag{145}$$

where ρ represents a scalar probability density, e^{ib} is a complex phase, and R is a rotor.

Comparatively, our wavefunction in $\mathcal{G}(\mathbb{R}^{3,1})$ is:

$$\phi = e^{-\frac{1}{4}\boldsymbol{\tau}\odot(a+\mathbf{x}+\mathbf{f}+\mathbf{v}+\mathbf{b})}\phi_0 \tag{146}$$

To approach David Hestenes' formulation of the wavefunction, it suffices to eliminate the terms $a \to 0$, $\mathbf{x} \to 0$ and $\mathbf{v} \to 0$, and to perform a substitution of the entries of the double-copy inner product (Equation 154), as follows:

$$\mathbf{w} \to \mathbf{u}^\ddagger \tag{147}$$

$$\mathbf{y} \to \mathbf{z}^\ddagger \tag{148}$$

As one of the copies is destroyed by the substitution, the double-copy inner product reduces to an inner product. Furthermore, with the elimination, the blade-3,4 conjugate is also reduced to the blade-4 conjugate, yielding

$$\langle \mathbf{u}, \mathbf{w}, \mathbf{y}, \mathbf{z} \rangle \to \langle \mathbf{u}, \mathbf{u}^\ddagger, \mathbf{z}^\ddagger, \mathbf{z} \rangle \cong \langle \mathbf{u}, \mathbf{z} \rangle = \sum_{i=1}^{m} \lfloor u_i^2 \rfloor_{2,4}(z_i^2) \tag{149}$$

Consequently, our wavefunction ϕ reduces to

$$\phi^2 = e^{-\frac{1}{2}\boldsymbol{\tau}\odot(\mathbf{f}+\mathbf{b})}\phi_0^2 \tag{150}$$

This shows that the 3+1D wavefunction (comprising a rotor $R(\tau) = e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}}$, a pseudo-scalar $e^{-\frac{1}{2}\tau_b\mathbf{b}}$ and a prior probability $\phi_0^2 = \sqrt{\rho}$) is a sub-structure of the general $\mathcal{G}(\mathbb{R}^{3,1})$ wavefunction. The primary difference is that our formulation also contains the generators of conservation laws and global symmetries, and that it lives in a blade 2-4 Hilbert space.

In this sub-structure, the observables are satisfied when

$$\lfloor \mathbf{O} \rfloor_{2,4} = \mathbf{O} \tag{151}$$

Let us now analyze the symmetry group of this wavefunction.

First, we note that the term $\mathbf{b}$ commutes with $\mathbf{f}$. They can be factored out as

$$e^{-\frac{1}{2}\boldsymbol{\tau}\odot(\mathbf{f}+\mathbf{b})}\phi_0^2 = e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{b}}e^{-\frac{1}{2}\tau_b\mathbf{f}}\phi_0^2 \tag{152}$$

Second, the term $\exp \mathbf{f}$ can be understood as the exponential map from the bivectors to the $\mathrm{Spin}_+(3,1)$ group and the term $\exp \mathbf{b}$ to $\mathrm{U}(1)$.

Finally, since $\mathrm{Spin}_+(3,1) \cap \exp \mathbf{b} = \{\pm 1\}$, it must be removed from the group product[11].

We conclude that the geometric components of the wavefunction corresponds to the following group

$$\mathrm{U}(1) \times (\mathrm{Spin}_+(3,1)/\{\pm 1\}) \cong \mathrm{Spin}^c(3,1) \tag{153}$$

4.12 Geometric Hilbert space in 3+1D (broken symmetry)

The substitution given by Equation 149 yields the following algebra of geometric observables:

Let $\mathbb{V}$ be an 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. $\forall \boldsymbol{\psi} \in \mathcal{A}(\mathbb{V})$, the inner product form

$$\begin{aligned} \langle \cdot, \cdot \rangle \quad : \quad & \mathbb{V} \times \mathbb{V} \longrightarrow \mathcal{G}(\mathbb{R}^{3,1}) \\ & \langle \mathbf{u}, \mathbf{w} \rangle \longmapsto \sum_{i=1}^{m} \lfloor u_i^2 \rfloor_{2,4} w_i^2 \end{aligned} \tag{154}$$

   is positive-definite when $\boldsymbol{\psi} \neq 0$; that is $\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle > 0$

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

$$\rho(\psi(q)) = \frac{1}{\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle} \det \psi(q), \tag{155}$$

   is either positive or equal to zero.

We note the following properties, features, and comments:

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

$$\sum_{\psi(q) \in \boldsymbol{\psi}} \rho(\psi(q)) = 1 \tag{156}$$

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

$$\langle\mathbf{T}\boldsymbol{\psi},\mathbf{T}\boldsymbol{\psi}\rangle = \langle\boldsymbol{\psi},\boldsymbol{\psi}\rangle \tag{157}$$

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

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

$$\langle\mathbf{O}\mathbf{u},\mathbf{w}\rangle = \langle\mathbf{u},\mathbf{O}\mathbf{w}\rangle \tag{158}$$

  is called an observable.

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

$$\langle\mathbf{O}\rangle = \frac{\langle\mathbf{O}\boldsymbol{\psi},\boldsymbol{\psi}\rangle}{\langle\boldsymbol{\psi},\boldsymbol{\psi}\rangle} \tag{159}$$

4.13 Gravity and electromagnetism in 3+1D

In 2D, we benefited from a coincidence of low dimensions, where the matrix representation of $\mathcal{G}(\mathbb{R}^2)$ was in $\mathbb{M}(2,\mathbb{R})$. As such, our wavefunction generated $\mathrm{GL}^+(2,\mathbb{R})$ which acted as the structure group of the frame bundle FX, and following a structure reduction from $\mathrm{GL}^+(2,\mathbb{R})$ to SO(2), a tetrad field was associated with the global section of the quotient bundle FX/SO(2) which led to a gravitized quantum theory.

In 4D, unlike in 2D where $\mathrm{SO}(2) = \mathrm{Spin}(2)$, the geometry of the wavefunction is not in SO but rather in Spin^c. And since Spin^c is not, in general, in GL^+, we cannot benefit from the same coincidences as in 2D.

Typically, to reach $\mathrm{Spin}(p,q)$ from the structure group $\mathrm{GL}(p+q)$, one would reduce $\mathrm{GL}(p+q)$ to $\mathrm{O}(p,q)$, then lift it to $\mathrm{Spin}(p,q)$. Here, however, we will use a different approach to get the spin connection.

Remarkably, 4D admits a coincidence that will allow us to embed the $\mathrm{Spin}^c(3,1)$ group into the $\mathrm{GL}^+(4,\mathbb{R})$ group, then take its quotient $\mathrm{FX}/\mathrm{Spin}^c(3,1)$ without having to lift; our solution already contains what is necessary to take this quotient.

The coincidence comes from the standard classification of real Clifford algebra[12] and from the fact that $\exp(\mathbf{f}+\mathbf{b}) \cong \mathrm{Spin}^c(3,1) \subset \exp \mathcal{G}(\mathbb{R}^{3,1})$. The diagram

$$\begin{array}{ccc} \mathcal{G}(\mathbb{R}^{3,1}) & \xrightarrow{f} & \mathbb{M}(4,\mathbb{R}) \\ \downarrow{\scriptstyle exp} & & \downarrow{\scriptstyle exp} \\ \exp \mathcal{G}(\mathbb{R}^{3,1}) & \xrightarrow{f} & \mathrm{GL}^+(4,\mathbb{R}) \end{array} \tag{160}$$

commutes by group homomorphisms. Since $\exp(\mathbf{f}+\mathbf{b}) \cong \mathrm{Spin}^c(3,1) \subset \exp \mathcal{G}(\mathbb{R}^{3,1})$, the map f embeds $\mathrm{Spin}^c(3,1)$ into $\mathrm{GL}^+(4,\mathbb{R})$. The inclusion of $\mathrm{Spin}^c(3,1)$ in $\exp \mathcal{G}(\mathbb{R}^{3,1})$ is required to break the symmetry into exactly a theory of gravity and of electromagnetism for charged fermions and into a $\mathrm{Spin}^c(3,1)$-valued quantum theory. We are now ready.

Let X^4 be a world manifold.

We first consider the tangent bundle TX along with its associated frame bundle FX. Our wavefunction acts on the frame bundle using the exponential map of multivectors $\exp \mathcal{G}(\mathbb{R}^{3,1}) \cong \exp \mathbb{M}(4,\mathbb{R})$ which generates $\mathrm{GL}^+(4,\mathbb{R})$.

The desired reduction is from $\exp \mathcal{G}(\mathbb{R}^{3,1})$ to the $\mathrm{Spin}^c(3,1)$ group. With its symmetry reduced, the wavefunction will assign an element of $\mathrm{Spin}^c(3,1)$ to each event $x \in X^4$. The connection that preserves the structure is a $\mathrm{Spin}^c(3,1)$ preserving connection. It relates to a theory of gravity and of electromagnetism for charged fermions. We note that since $\mathrm{SO}(3,1)\times\mathrm{U}(1)$ is a quotient $\mathrm{Spin}^c(3,1)$, the cosets are further associable with the inner products. Thus, the global section of the quotient bundle $\mathrm{FX}/\mathrm{SO}(3,1)$ associates with a tetrad field that uniquely determines a pseudo-Riemannian metric. As for the $\mathrm{U}(1)$-bundle, it is simply the geometric setting for electromagnetism. Finally, the frame bundle is a natural bundle that admits general covariant transformations, which are the symmetries of the gravitation theory on X^4[9]. This is the geometric setting for gravity.

4.14 Metric interference in 3+1D

A geometric wavefunction would allow a larger class of interference patterns than complex interference. The geometric interference pattern includes the ways in which the geometry of a probability measure can interfere constructively or destructively and includes interference from rotations, phases, boosts, shears, spins, and dilations.

In the case of 4D *metric interference* (shown below), the interference pattern is associated with a superposition of elements of the group $\mathrm{Spin}^c(3,1)$, whose subgroup $\mathrm{SO}(3,1)$ associates to a superposition of inner products in the quotient.

It is possible that a *sensitive* Aharonov–Bohm effect experiment on gravity[13] could detect special cases of the geometric phase and interference 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:

$$\det(\mathbf{u}+\mathbf{v}) = \det\mathbf{u} + \det\mathbf{v} + \mathbf{u}\cdot\mathbf{v} \tag{161}$$

The determinants $\det\mathbf{u}$ and $\det\mathbf{v}$ are a sum of probabilities, whereas the dot product term $\mathbf{u}\cdot\mathbf{v}$ represents the interference term.

Such can be obtained following a transformation of a wavefunction $|\psi\rangle = \begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix}$ such that the multivectors are mapped to a linear combination of two multivectors:

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{u}+\mathbf{v}\\ \mathbf{u}-\mathbf{v}\end{bmatrix} \tag{162}$$

The dot product defines a bilinear form.

$$\cdot \quad : \qquad \mathcal{G}(\mathbb{R}^{m,n})\times\mathcal{G}(\mathbb{R}^{m,n}) \longrightarrow \mathbb{R} \tag{163}$$

$$\mathbf{u}\cdot\mathbf{v} \longmapsto \frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det\mathbf{u} - \det\mathbf{v}) \tag{164}$$

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 has this form

$$\begin{aligned}
&\frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det\mathbf{u} - \det\mathbf{v}) && (165)\\
&= \frac{1}{2}\left((\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v}) - \mathbf{u}^\ddagger\mathbf{u} - \mathbf{v}^\ddagger\mathbf{v}\right) && (166)\\
&= \mathbf{u}^\ddagger\mathbf{u} + \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} + \mathbf{v}^\ddagger\mathbf{v} - \mathbf{u}^\ddagger\mathbf{u} - \mathbf{v}^\ddagger\mathbf{v} && (167)\\
&= \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} && (168)
\end{aligned}$$

In 3+1D, it has this form.

$$\begin{aligned}
&\frac{1}{2}(\det(\mathbf{u}+\mathbf{v}) - \det\mathbf{u} - \det\mathbf{v}) && (169)\\
&= \frac{1}{2}\left(\lfloor(\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v})\rfloor_{3,4}(\mathbf{u}+\mathbf{v})^\ddagger(\mathbf{u}+\mathbf{v}) - \lfloor\mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} - \lfloor\mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v}\right) && (170)\\
&= \frac{1}{2}\left(\lfloor\mathbf{u}^\ddagger\mathbf{u} + \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} + \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}(\mathbf{u}^\ddagger\mathbf{u} + \mathbf{u}^\ddagger\mathbf{v} + \mathbf{v}^\ddagger\mathbf{u} + \mathbf{v}^\ddagger\mathbf{v}) - \ldots\right) && (171)
\end{aligned}$$

$$
\begin{aligned}
&= \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\
&\quad + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\
&\quad + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\
&\quad + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} - \dots \qquad (172)
\end{aligned}
$$

$$
\begin{aligned}
&= \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\
&\quad + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{u}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\
&\quad + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{u}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{v} \\
&\quad + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{u} + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{u}^\ddagger\mathbf{v} + \lfloor \mathbf{v}^\ddagger\mathbf{v}\rfloor_{3,4}\mathbf{v}^\ddagger\mathbf{u} \qquad (173)
\end{aligned}
$$

We now consider simpler interference patterns.

Interference in 3+1D:

As seen previously, the substituted double-copy inner product reduces to an inner product (Equation 149). The interference pattern[14] is given as follows:

$$
\begin{aligned}
\det(\mathbf{u}+\mathbf{v}) &= \lfloor \mathbf{u}+\mathbf{v}\rfloor_{2,4}(\mathbf{u}+\mathbf{v}) && (174)\\
&= \lfloor \mathbf{u}\rfloor_{2,4}(\mathbf{u}+\mathbf{v}) + \lfloor \mathbf{v}\rfloor_{2,4}(\mathbf{u}+\mathbf{v}) && (175)\\
&= \lfloor \mathbf{u}\rfloor_{2,4}\mathbf{u} + \lfloor \mathbf{u}\rfloor_{2,4}\mathbf{v} + \lfloor \mathbf{v}\rfloor_{2,4}\mathbf{u} + \lfloor \mathbf{v}\rfloor_{2,4}\mathbf{v} && (176)\\
&= \det\mathbf{u} + \det\mathbf{v} + \lfloor \mathbf{u}\rfloor_{2,4}\mathbf{v} + \lfloor \mathbf{v}\rfloor_{2,4}\mathbf{u} && (177)
\end{aligned}
$$

Now replacing $\mathbf{u} = \rho_u e^{-\frac{1}{2}\tau_b\mathbf{b}_u} e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}_u}$ and $\mathbf{v} = \rho_v e^{-\frac{1}{2}\tau_b\mathbf{b}_v} e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}_v}$

$$
= |\rho_u|^2 + |\rho_v|^2 + \rho_u\rho_v\left(e^{\frac{1}{2}\tau_b\mathbf{b}_u} e^{\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}_u} e^{-\frac{1}{2}\tau_b\mathbf{b}_v} e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}_v} + e^{\frac{1}{2}\tau_b\mathbf{b}_v} e^{\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}_v} e^{-\frac{1}{2}\tau_b\mathbf{b}_u} e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}_u}\right) \qquad (178)
$$

Due to the presence of $\mathbf{f}$ and $\mathbf{b}$, the geometric richness of the interference pattern exceeds that of the 2D case. The term $\mathbf{f}$ associates with a non-commutative interference effect in the interference pattern, which distinguishes it from (the entirely commutative) complex interference and could presumably be identified experimentally in a properly constructed interference experiment.

4.15 Dirac current

David Hestenes[10] defines the Dirac current in the language of geometric algebra as:

$$
\mathbf{j} = \psi^\ddagger\gamma_0\psi = \rho R^\ddagger\gamma_0 R = \rho e_0 = \rho v \qquad (179)
$$

where v is the proper velocity.

In our formulation, this relation also holds; the Dirac current represents the action of the wavefunction on the unit timelike vector in the tangent space on

X^4. Specifically, the Dirac current is a statistically weighted Lorentz action on γ_0:

$$
\begin{aligned}
\mathbf{j} &= \psi^{\ddagger}\gamma_0\psi && (180)\\
&= e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}+\frac{1}{2}\tau_b\mathbf{b}}\phi_0\gamma_0 e^{\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}+\frac{1}{2}\tau_b\mathbf{b}}\phi_0 && (181)\\
&= \phi_0^2 e^{-\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}}\gamma_0 e^{\frac{1}{2}\boldsymbol{\tau}\odot\mathbf{f}} && (182)\\
&= \rho(\boldsymbol{\tau})e_0(\boldsymbol{\tau}) && (183)\\
&= \rho(\boldsymbol{\tau})v(\boldsymbol{\tau}) && (184)
\end{aligned}
$$

We now have all the tools required to construct particle physics by exhausting the remaining geometry of our solution.

4.16 SU(2) × U(1) group

Our wavefunction transforms as a group under multiplication. We now ask, what is the most general multivector $e^{\mathbf{u}}$ which leaves the Dirac current invariant?

$$\psi^{\ddagger}(e^{\mathbf{u}})^{\ddagger}\gamma_0 e^{\mathbf{u}}\psi = \psi^{\ddagger}\gamma_0\psi \iff (e^{\mathbf{u}})^{\ddagger}\gamma_0 e^{\mathbf{u}} = \gamma_0 \quad (185)$$

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. Finally, the base $\gamma_0\gamma_1\gamma_2\gamma_3$ anti-commutes with γ_0 and 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

$$e^{\mathbf{u}} = \exp\left(\frac{1}{2}F_{12}\gamma_1\gamma_2 + \frac{1}{2}F_{13}\gamma_1\gamma_3 + \frac{1}{2}F_{23}\gamma_2\gamma_3 + \frac{1}{2}\mathbf{b}\right) \quad (186)$$

We can rewrite the bivector basis with the Pauli matrices

$$
\begin{aligned}
\gamma_2\gamma_3 &= \mathbf{i}\sigma_x && (187)\\
\gamma_1\gamma_3 &= \mathbf{i}\sigma_y && (188)\\
\gamma_1\gamma_2 &= \mathbf{i}\sigma_z && (189)\\
\mathbf{b} &= \mathbf{i}b && (190)
\end{aligned}
$$

After replacements, we obtain

$$e^{\mathbf{u}} = \exp\frac{1}{2}\mathbf{i}(F_{12}\sigma_z + F_{13}\sigma_y + F_{23}\sigma_x + b) \quad (191)$$

The terms $F_{23}\sigma_x + F_{13}\sigma_y + F_{12}\sigma_z$ and b are responsible for SU(2) and U(1) symmetries, respectively[15, 16].

4.17 SU(3) group

The invariance transformation identified by the 3+1D algebra of geometric observables (Equation 157) are $\mathbf{T}^\ddagger\mathbf{T}=\mathbf{I}$, $\mathbf{T}^\dagger\mathbf{T}=\mathbf{I}$ and $\lfloor\mathbf{T}\rfloor_{2,4}\mathbf{T}=\mathbf{I}$. In the first case, the identified evolution is bivectorial rather than unitary.

As we did for the SU(2) × U(1) case, we ask, in this case, what is the most general bivectorial evolution that leaves the Dirac current invariant?

$$\mathbf{f}^\ddagger\gamma_0\mathbf{f}=\gamma_0 \tag{192}$$

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

$$\mathbf{f}=F_{01}\gamma_0\gamma_1+F_{02}\gamma_0\gamma_2+F_{03}\gamma_0\gamma_3+F_{23}\gamma_2\gamma_3+F_{13}\gamma_1\gamma_3+F_{12}\gamma_1\gamma_2 \tag{193}$$

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

$$\begin{aligned}\mathbf{f}^\ddagger\gamma_0\mathbf{f}=-\mathbf{f}\gamma_0\mathbf{f}=&(F_{01}^2+F_{02}^2+F_{03}^2+F_{13}^2+F_{23}^2+F_{12}^2)\gamma_0 &(194)\\ &+(-2F_{02}F_{12}+2F_{03}F_{13})\gamma_1 &(195)\\ &+(-2F_{01}F_{12}+2F_{03}F_{23})\gamma_2 &(196)\\ &+(-2F_{01}F_{13}+2F_{02}F_{23})\gamma_3 &(197)\end{aligned}$$

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

$$\begin{aligned}-2F_{02}F_{12}+2F_{03}F_{13}&=0 &(198)\\ -2F_{01}F_{12}+2F_{03}F_{23}&=0 &(199)\\ -2F_{01}F_{13}+2F_{02}F_{23}&=0 &(200)\end{aligned}$$

leaving only

$$\mathbf{f}^\ddagger\gamma_0\mathbf{f}=(F_{01}^2+F_{02}^2+F_{03}^2+F_{13}^2+F_{23}^2+F_{12}^2)\gamma_0. \tag{201}$$

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:

$$\mathbf{f}=(F_{01}+\mathbf{i}F_{23})\gamma_0\gamma_1+(F_{02}+\mathbf{i}F_{13})\gamma_0\gamma_2+(F_{03}+\mathbf{i}F_{12})\gamma_0\gamma_3 \tag{202}$$

Then, with the nullification of the cross-product and equating $F_{01}^2+F_{02}^2+F_{03}^2+F_{13}^2+F_{23}^2+F_{12}^2$ to unity, we can understand the bivectorial evolution when constrained by the Dirac current to be a realization of the SU(3) group[16].

4.18 Satisfiability of geometric observables in 4D

In 4D, an observable must satisfy equation 143. Let us now verify that geometric observables are satisfiable in 4D. For simplicity, let us take m in equation 154 to be 1. Then,

$$\lfloor (\mathbf{O}u)^{\ddagger} w \rfloor_{3,4} y^{\ddagger} z = \lfloor u^{\ddagger} \mathbf{O} w \rfloor_{3,4} y^{\ddagger} z = \lfloor u^{\ddagger} w \rfloor_{3,4} (\mathbf{O}y)^{\ddagger} z = \lfloor u^{\ddagger} w \rfloor_{3,4} y^{\ddagger} \mathbf{O} z \tag{203}$$

where u_1, w_1, y_1 and z_1 are multivectors.

Let us investigate.

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

We thus find that observables are satisfiable in the general 4D case. We also recall that in 3+1D, the observable reduces to $\lfloor \mathbf{O} \rfloor_{2,4} = \mathbf{O}$, which is also satisfiable.

4.19 Unsatisfiability of geometric observables in 6D and above

At dimensions of 6 or above, the corresponding observable relation cannot be satisfied. To see why, we look at the results[17] of Acus et al. regarding the 6D multivector norm. The authors performed an exhaustive computer-assisted search for the geometric algebra expression for the determinant in 6D; as conjectured, they found no norm defined via self-products. The norm is a linear combination of self-products.

The system of linear equations is too long to list in its entirety; the author gives this mockup:

$$a_0^4 - 2a_0^2 a_{47}^2 + b_2 a_0^2 a_{47}^2 p_{412} p_{422} + \langle 72 \text{ monomials} \rangle = 0 \tag{204}$$

$$b_1 a_0^3 a_{52} + 2b_2 a_0 a_{47}^2 a_{52} p_{412} p_{422} p_{432} p_{442} p_{452} + \langle 72 \text{ monomials} \rangle = 0 \tag{205}$$

$$\langle 74 \text{ monomials} \rangle = 0 \tag{206}$$

$$\langle 74 \text{ monomials} \rangle = 0 \tag{207}$$

The author then produces the special case of this norm that holds only for a 6D multivector comprising a scalar and grade 4 element:

$$s(B) = b_1 B f_5(f_4(B) f_3(f_2(B) f_1(B))) + b_2 B g_5(g_4(B) g_3(g_2(B) g_1(B))) \tag{208}$$

Even in this simplified special case, formulating a linear relationship for observables is doomed to fail. Indeed, the real portion of the observable cannot be extracted from the equation. We find that for any function f_i and g_i, the coefficient b_1 and b_2 will frustrate the equality:

$$b_1\mathbf{O}Bf_5(f_4(B)f_3(f_2(B)f_1(B))) + b_2Bg_5(g_4(B)g_3(g_2(B)g_1(B))) \tag{209}$$

$$= b_1Bf_5(f_4(B)f_3(f_2(B)f_1(B))) + b_2\mathbf{O}Bg_5(g_4(B)g_3(g_2(B)g_1(B))) \tag{210}$$

Equations 209 and 210 can only be equal if $b_1 = b_2$; however, the norm $s(B)$ requires both to be different. Consequently, the relation for observables in 6D is unsatisfiable even by real numbers.

Thus, in our solution, observables are satisfiable in 6D.

Furthermore, since the norms involve more sophisticated systems of linear equations in higher dimensions, this result is likely to generalize to all dimensions above 6.

4.20 Defective probability measure in 3D and 5D

The 3D and 5D cases (and possibly all odd-dimensional cases of higher dimensions) contain a number of irregularities that make them defective to use in this framework. Let us investigate.

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

$$\mathbf{u} = a + x\sigma_x + y\sigma_y + z\sigma_z + q\sigma_y\sigma_z + v\sigma_x\sigma_z + w\sigma_x\sigma_y + b\sigma_x\sigma_y\sigma_z \tag{211}$$

is

$$\mathbf{u} \cong \begin{bmatrix} a + ib + iw + z & iq - v + x - iy \\ iq + v + x + iy & a + ib - iw - z \end{bmatrix} \tag{212}$$

and the determinant is

$$\det \mathbf{u} = a^2 - b^2 + q^2 + v^2 + w^2 - x^2 - y^2 - z^2 + 2i(ab - qx + vy - wz) \tag{213}$$

The result is a complex-valued probability. Since a probability must be real-valued, the 3D case is defective in our solution and cannot be used. In theory, it can be fixed by defining a complex norm to apply to the determinant:

$$\langle \mathbf{u}, \mathbf{u} \rangle = (\det \mathbf{u})^\dagger \det \mathbf{u} \tag{214}$$

However, defining such a norm would entail a double-copy inner product of 4 multivectors, but the space is only 3D, not 4D (so why four?). It would also break the relationship between trace and probability that justified its usage in statistical mechanics.

Consequently, this case appears to us to be defective.

Perhaps, instead of $\mathcal{G}(\mathbb{R}^3)$ multivectors, we ought to use 3×3 matrices in 3D? Alas, 3×3 matrices do not admit a geometric algebra representation because they are not isomorphic with $\mathcal{G}(\mathbb{R}^3)$. Indeed, $\mathcal{G}(\mathbb{R}^3)$ has 8 parameters and 3×3 matrices have 9. 3×3 matrices are not representable geometrically in the same sense that 2×2 matrices are with $\mathcal{G}(\mathbb{R}^2)$.

In $\mathcal{G}(\mathbb{R}^{4,1})$, the algebra is isomorphic to complex 4×4 matrices. In this case, the determinant and probability would be complex-valued, making the case defective. Furthermore, 5×5 matrices have 25 parameters, but $\mathcal{G}(\mathbb{R}^{4,1})$ multivectors have 32 parameters.

4.21 The dimensions that admit observable geometry

Our solution is non-defective in the following dimensions:

- $\mathbb{R}$: This case corresponds to familiar 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 $\rho(q) = \frac{1}{Z(\beta)} \exp\big(-\beta E(q)\big)$.

- $\mathbb{C} \cong \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$: This case corresponds to familiar non-relativistic quantum mechanics.

However, neither of these cases contain geometry. The only case that contain observable geometry are:

- $\mathcal{G}(\mathbb{R}^2)$: This case corresponds to the geometric quantum theory in 2D. Its $\mathrm{GL}^+(2)$ symmetry breaks into a theory of gravity FX/SO(2) and into a quantum theory valued in SO(2).

- $\mathcal{G}(\mathbb{R}^{3,1})$: This case is valid. Like the 2D case, it also corresponds to a geometric quantum theory. As such, its symmetry will break into a theory of gravity and a relativistic wavefunction. But unlike the 2D case, the wavefunction further admits an invariance with respect to the $\mathrm{SU}(2)\times\mathrm{U}(1)$ and SU(3) gauge groups.

In contrast, our solution is defective in the following dimensions:

- $\mathcal{G}(\mathbb{R}^3)$: In this case, the probability measure is complex-valued.

- $\mathcal{G}(\mathbb{R}^{4,1})$: In this case, the probability measure is complex-valued.

- 6D and above: For $\mathcal{G}(\mathbb{R}^n)$, where $n \geq 6$, no observables satisfy the corresponding observable equation, in general.

We may thus say that 3D and 5D fail to normalize, and 6D and above fail to satisfy observables. Consequently, in the general case of our solution, normalizable geometric observables cannot be satisfied beyond 4D. This suggests an intrinsic limit to the dimensionality of observable geometry, and by extension to spacetime.

5 Discussion

Our goal was to devise a framework in which the laws of physics can be derived rather than postulated. Furthermore, we wish to derive them such that in some sense the result is optimal.

One way to arrive at the laws of physics is to inspect the scientific literature, then intuit a number of axioms, and finally submit them to experimental testing for falsification. Such methodology can demonstrate scientific fitness (or more precisely, lack thereof), but cannot demonstrate optimality. Indeed, as pure mathematics is unaware of the scientific fitness of any sets of axioms, it cannot assign them a score apriori. Furthermore, such methodology clashes with the logical definition of an axiom because axioms in logic are held true by definition, but laws in physics are entailed by measurement.

Our proposal capitalizes on a missed opportunity to constrain the set of all theories by the structure of all measurements, thus constraining the theories to be empirical. Under this constraint, physics is found as the solution that maximizes the information associated with the construction by the observer of a message of realized measurements. The optimization problem assigns a score (measured in terms of quantity of information) to each possible empirical theory and selects the maximally informative one for a given measurement structure. Physics is thus the provable explanatory maximum for realized measurements.

Defining the problem is this manner requires a single axiom, the *fundamental structure of measurements* (Axiom 1) and such is sufficient to entail the *fundamental theorem of physics* (Theorem 1) as its main result. The fundamental structure of measurements is a mathematical expression motivated solely as a best empirical fit to the structure of the measurements that are found in nature.

With this foundation, the pervasive platonic defect induced by defining laws as axioms, rather than deriving them from the measurements that entail them, is now corrected. In our formulation, and for the first time, the foundation of physics is completely consistent with physics being an empirical science because it refers exclusively to the structure of measurements that empirically fit nature, and lacks any other importations.

The techniques of statistical mechanics are used abundantly in our work. The complete correspondence between an ordinary system of statistical mechanics and ours is as follows.

Table 1: Correspondence

Concept	Energy Constraint	Measurement Constraint
Entropy	Boltzmann	Shannon
Measure	Gibbs	Born rule
Constraint	Energy meter	Phase-invariant instrument
Micro-state	Energy values	Measurement results
Lagrange multiplier	Temperature	Entropic flow
Ontology	Ergodic system	Construction of a message

In the correspondence, using the Shannon entropy instead of the Boltzmann entropy changes the ontology from ergodic systems to the construction of a message (in the sense of the communication theory of Claude Shannon[18]) of measurements. The construction of such a message by an observer carries information; it is associated to the registration of a "click"[19] on a screen or an incidence counter. Since the message is received by the observer, we are not dealing with entropy but information, even though the equations are similar to those of statistical physics.

The correspondence is consistent with the general intuition that *random information* must be axiomatic, as, by definition, it cannot be derived from any earlier principles. Ultimately, it is preferable to consider *the message of measurements* (whose random elements associates to information), rather than the wavefunction, to be the axiomatic foundation of the theory. As shown, the latter can be derived from the former but not vice versa (the measurement collapse problem prevents the mathematical derivation of the elements of the message).

The probabilistic interpretation of the wavefunction via the Born rule is inherited from the solution, and is explained as the measure that maximizes the information of measurements. The wavefunction is derived and consequently is not considered axiomatic. The measurement postulate is also not axiomatic, because the resulting probability measure is parametrized over the possible measurement values of the system (and consequently, as per definition of a probability measure in statistical mechanics, the elements of the parametrization already corresponds to the possible measurements). Even the appropriate Hilbert space is mechanically recovered by the solution and need not be added by hand. The optimization problem's ability to mechanically produce the correct Hilbert space is its main practical advantage. This allows the seamless extension of the Hilbert space to accommodate the structure of general linear measurements. It turns out that in 2D a *geometric* Hilbert space is required to accommodate general linear measurements, and in 3+1D a *double-copy* geometric Hilbert space is required to accommodate the same. Both the 2D and the 3+1D geometric Hilbert spaces contain gravity, but the second one also contain the standard model. Both spaces are mathematically well-defined, but are highly non-obvious without the benefit of our optimization problem to mechanically derive them.

6 Conclusion

We proposed to maximize the information associated with a construction by the observer of a message of measurement under the constraint of a geometric measurement apparatus. The resulting probability measure supports a geometry richer than what could previously be supported in either statistical physics or quantum mechanics. Accommodating all possible geometric measurements entails a geometric wavefunction, for which the Born rule is extended to the determinant. This substantially extends the opportunity to capture all fundamental physics within a single framework. The framework produces solutions for 2D and

4D in which general observables are normalizable. 4D stands out as the largest geometry that satisfies the conditions for having normalizable observables in the general case. A gravitized standard model results from the frame bundle FX of a world manifold, whose structure group is generated by $\exp \mathcal{G}(\mathbb{R}^{3,1})$ (which is isomorphic to $\exp \mathbb{M}(4, \mathbb{R})$ and as such generates to $\mathrm{GL}^+(4, \mathbb{R})$ up to isomorphism), undergoing symmetry breaking to $\mathrm{Spin}^c(3,1)$. The global sections of the quotient bundle $\mathrm{FX}/\mathrm{SO}(3,1)$ identify with a pseudo-Riemannian metric and the natural bundles to general covariant transformations. The connection is a Spin^c-preserving connection. The group $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$ is recovered in the broken symmetry and associates to the invariant transformations under the action of the wavefunction on a unit timelike vector of the tangent space, and preserving the Dirac current. Finally, an interpretation of quantum mechanics, i.e., the maximally-informative interpretation, is proposed; the structure of measurements 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 the probability measure, maximizing the information of this message, is the geometric wavefunction accompanied by the geometric Born rule. It is the only interpretation whose mathematical formulation is sufficiently precise to recover, by itself, the full machinery of quantum physics (and even improve upon it). Finally, as the solution to an optimization problem on information, we concluded that physics, distilled to its conceptually simplest expression, is the solution that provably makes realized measurements maximally informative to the observer. Physics is the provable explanatory maximum for realized measurements.

7 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] Edwin T Jaynes. Information theory and statistical mechanics (notes by the lecturer). *Statistical physics 3*, page 181, 1963.

[6] Edwin T Jaynes. Prior probabilities. *IEEE Transactions on systems science and cybernetics*, 4(3):227–241, 1968.

[7] Marc Lachieze-Rey. Connections and frame bundle reductions. *arXiv preprint arXiv:2002.01410*, 2020.

[8] ET Tomboulis. General relativity as the effective theory of g l (4, r) spontaneous symmetry breaking. *Physical Review D*, 84(8):084018, 2011.

[9] G Sardanashvily. Classical gauge gravitation theory. *International Journal of Geometric Methods in Modern Physics*, 8(08):1869–1895, 2011.

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

[11] Nicholas Todoroff (https://math.stackexchange.com/users/1068683/nicholas todoroff). Does the exponential of a bivector + pseudo-scalar in cl(3,1) maps to $\mathrm{spin}^c(3,1)$? Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/4659083 (version: 2023-03-14).

[12] Nicholas Todoroff (https://math.stackexchange.com/users/1068683/nicholas todoroff). Does $\exp \mathbf{u}$ where $\mathbf{u} \in \mathrm{CL}(3,1)$ maps to $\mathrm{GL}^+(4,\mathbb{R})$ or its double cover? Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/4665079 (version: 2023-03-23).

[13] 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.

[14] Bohdan I Lev. Wave function as geometric entity. *Journal of Modern Physics*, 3:709–713, 2012.

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

[16] 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.

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

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

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