

The Arithmetic of Uncertainty unifies Quantum Formalism and Relativistic Spacetime

John Skilling,¹ Kevin H. Knuth,^{2*}

¹Maximum Entropy Data Consultants Ltd, Kenmare, Ireland

²University at Albany (SUNY) Albany NY 12222, USA

*Corresponding author; E-mail: kknuth@albany.edu.

Abstract

The theories of quantum mechanics and relativity dramatically altered our understanding of the universe ushering in the era of modern physics. Quantum theory deals with objects probabilistically at small scales, whereas relativity deals classically with motion in space and time.

We show here that the mathematical structures of quantum theory and of relativity follow together from pure thought, defined and uniquely constrained by the same elementary “combining and sequencing” symmetries that underlie standard arithmetic and probability. The key is uncertainty, which inevitably accompanies observation of quantity and imposes the use of *pairs* of numbers.

The symmetries then lead directly to the use of complex “ $\sqrt{-1}$ ” arithmetic, the standard calculus of quantum mechanics, and the Lorentz transformations of relativistic spacetime. One dimension of time and three dimensions of space are thus *derived* as the profound and inevitable framework of physics.

1 Introduction

Science relies on numeric quantification, which can be traced back to Euclid, Galileo, and Newton. While many now take this success for granted, some scientists have pondered the “unreasonable effectiveness of mathematics” [1, 2] as well as its necessity [3], and others continue to strive to understand the foundations of the quantum mechanical formalism [4, 5, 6, 7, 8, 9]. Such questions are deep. For example, we quantify speed with a single real number, a scalar.

But we must quantify velocity with three real numbers. Why? And how do we know?

This paper is about engineering the formalism of quantification [10]. To accomplish this, we use the basic symmetries beneath combination and sequencing, which underpin formal science. Only such supremely elementary assumptions have such sufficient range and authority to be an acceptable base for the fundamental language of wide-ranging science. Requiring some subtle detail (continuity perhaps) would be less compelling because it could more easily be denied. Robust foundation needs to be simple. Hoping for wide accessibility, our account eschews excessive detail and is intended to be accessible to neophyte undergraduate students of reasonable diligence and moderate numeracy.

Classical objects might optimistically be measurable to arbitrary precision: “*Given any precision requirement, there could exist technology to accomplish it*”. That would mean that quantity could be treated as a single number, ignoring any intrinsic but negligible uncertainty. But this is a double limit and the pessimistic view “*Given any technological ability, there will always exist requirements that defeat it*” could also be held. Classically, it’s not clear whether “*requirement* $\rightarrow 0$ ” or “*technology* $\rightarrow 0$ ” should dominate, which leaves the status of single-number quantities uncomfortably in question.

Fortunately there is no ambiguity in the quantum world, where an elementary target can not be measured with precision by some arbitrarily smaller probe because there are no such probes. It becomes impossible to bootstrap complete knowledge of a target because we start correspondingly ignorant about any

probe that we might use. It follows that a faithful description of a quantum target requires a *pair* of numbers representing a fusion of quantity and uncertainty. The connection will be more intimate than just “quantity \pm error bar”, which would really be just a conventional couple of scalars.

The basic symmetries impose a specific calculus on number pairs, in keeping with but more subtle than standard scalar arithmetic. We derive *complex arithmetic* operating on pairs which we recognise as quantum amplitudes, observable through modulus-squared probabilities. Not only do we construct the Feynman picture of quantum mechanics, but we find that these same symmetries also lead to the Pauli matrices which generate spin, energy and momentum, and beyond that to 3+1-dimensional relativistic spacetime. The physics of quantity-with-uncertainty is to be described within this required mathematical formalism.

This adopts the strategy of [11] to develop mathematical language in accordance with the relevant fundamental symmetries (here, of physics). With this perspective, it is not surprising that mathematics works as the language of physics [12]. The mathematics we use works because it is engineered to work.

Our guiding principle is parsimony of laws. Any operation that our mathematics does allow should be allowed unless prevented by some new law, and that means that the mathematics must be as simple as we can make it. If we occasionally stray into technical language, it is not with intent to claim erudition or sophistication, but rather to reach out to those of greater erudition and sophistication than ourselves while aiming to explain to wider readership why the framework of physics must be as it is.

2 Addition and Multiplication

Sum and product rules are the foundation of arithmetic and thence of the rich structure of mathematics that science uses to model the physical world.

We assume that separate objects can exist. Although we illustrate this here with spots on the faces of dice, we do not attempt to define the nature of the objects in question. Applications are legion and we do not place limits on the objects or properties that users might have in mind. We just assume **commutativity** (order doesn't matter for the purpose in hand) illustrated with

$$\underbrace{\boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} \text{ with } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} = \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \text{ with } \boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} \quad (1)$$

commutative, $a+b = b+a$

and **associativity** (brackets don't matter either)

$$\underbrace{\left(\boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} \text{ with } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \right) \text{ with } \boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} = \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \text{ with } \left(\boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \text{ with } \boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} \right) \quad (2)$$

associative, $(a+b)+c = a+(b+c)$

This associative commutativity implies that the mathematical representation of quantification is additive [13, 14, 15] (up to isomorphism, which allows changing the labels while preserving the content). Usage of addition is the *sum rule*, here obtained in a way that will upgrade into quantum theory.

We also assume that what can be added up can be replicated, subject to left- and right-**distributivity** (replication applies to any target)

$$\underbrace{4 \text{ of } \left(\boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} \text{ with } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \right) = \left(4 \text{ of } \boxed{\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix}}} \right) \text{ with } \left(4 \text{ of } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \right) \\ (3 \text{ with } 5) \text{ of } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} = \left(3 \text{ of } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \right) \text{ with } \left(5 \text{ of } \boxed{\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}}} \right)} \quad (3)$$

distributive, $a(b+c) = ab+ac$ and $(a+b)c = ac+bc$

and non-trivial **associativity** (replications can be chained without forced an-

nihilation)

$$\underbrace{3 \text{ of } (4 \text{ of } 2 \text{ of } \boxed{\bullet}) = (3 \text{ of } 4 \text{ of } 2 \text{ of } \boxed{\bullet})}_{\text{associative, } a(bc) = (ab)c} \quad (4)$$

This associative distributivity implies that its representation is multiplicative (up to change of units, for consistency with addition). Partitioning is the inverse, where proportions multiply down instead of replicates multiplying up [16, 17, 18, 19, 15]. Usage of multiplication is the *product rule*.

Associative commutativity and associative distributivity are our *foundational symmetries*.

Nothing else is needed. Children use these informally as they learn about addition and multiplication through shuffling and grouping. The supreme simplicity of these ideas tokens the generality of application that we need for the deepest basis of science. This explains the success of mathematical modelling of a world in which separate objects can exist (commutativity) and which can behave independently of others (distributivity). The symmetries force arithmetical rules to which we can have no alternative. The content of our modelling is up to us, but the language (in this case standard arithmetic) is defined.

One can think of the mathematics used in science as having been engineered to be consistent with these symmetries, thus ensuring its success. It is then no mystery that mathematics works [1, 2], because it could not have been any other way [12]. The uniqueness (up to isomorphism [19, 15, 20]) and extreme familiarity of the rules appear to give them independent status, whereas in fact they were deliberately engineered. **If the foundational symmetries are accepted, the rules are forced, and the resulting mathematics becomes**

the quantitative language of physics.

Specifically, measure theory applies the sum rule to quantification, with additivity being the unique formalism for ubiquitous situations. Probability calculus applies the sum and product rules to partitioning of allowed possibilities [15, 20], which allows us to learn about the world by eliminating some of what was previously deemed possible. Again, probability is the unique calculus in ubiquitous situations. The content of the calculus is up to us, but the language is forced.

There are many applications of the rules, in human affairs as well as in science. For example, money is an application of measure leading to betting as a subsidiary application of probability [21, 22]. But applications are not foundations.

Where objects obeying these symmetries have only one relevant property, the standard representation is scalar. However, objects may have several properties. For example, six-sided dice have individuality (1 per die) and can display different numbers of spots (ranging from 1 to 6). Those properties, the number of dice thrown and the number of displayed dots, are separately additive.

Multidimensional addition is straightforwardly componentwise. Multidimensional multiplication, though, is not quite uniquely defined by the founding symmetries. Distributivity requires that a product is bilinear (linear in each factor) but associativity does not altogether remove the remaining ambiguity. Investigating this leads simply and directly to the basic language of physics in the form of quantum formalism and relativistic spacetime.

3 Quantum Foundation

Our world is quantised. There is an irreducible quantum of targets of a given type, and there are likewise irreducible investigative probes. Our knowledge always derives from interactions, which means that we can never attain complete knowledge of any individual object, whether it be a target investigated with an incompletely known probe, or a probe interacting with an incompletely known target. Consequently, our knowledge of objects will always be accompanied by inherent uncertainty. The connection between quantity and uncertainty is potentially more intimate than just “quantity \pm error bar”, which means that **our description needs to fuse quantity with uncertainty into two-parameter “pairs”** that we write as

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \dots \quad (5)$$

Precisely how quantity and uncertainty are to be encoded by these pairs is to be determined.

Associative commutativity

$$\underbrace{\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}}_{\text{commutative}} \quad \text{and} \quad \underbrace{(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})}_{\text{associative}} \quad (6)$$

implies the sum rule

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \quad (7)$$

now in two-dimensional pair-wise form.

Upon interaction, associative distributivity then gives

$$\underbrace{\begin{aligned} \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) &= \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c} \\ (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} &= \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} \end{aligned}}_{\text{distributive, left and right}} \quad \text{and} \quad \underbrace{(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c})}_{\text{associative}} \quad (8)$$

because summation has to remain linear regardless of probing or targeting context (distributivity) and probing or targeting is a sequential process (associative). Left distributivity ensures that a pair product is linear in the second factor and right distributivity ensures linearity in the first, so that the multiplication is bilinear, taking the form

$$(\mathbf{x} \cdot \mathbf{y})_i = \sum_{j,k=1}^2 \gamma_{ijk} x_j y_k \quad (9)$$

where the γ 's are eight constants which take standard values in standardised coordinates. **This is why physics is fundamentally linear.** Probing with a pair \mathbf{x} applies a linear 2×2 matrix

$$(\mathbf{x} \cdot)_ik = \sum_{j=1}^2 \gamma_{ijk} x_j \quad (10)$$

to the target pair \mathbf{y} .

Unlike for scalars, pair multiplication is not unique. We encounter similarly nonunique multiplication through dot and cross products in vector calculus (though those products are not full-rank 3-vectors). Here, we seek multiplication in which the $\mathbb{R}^2 \times \mathbb{R}^2$ pair products remain full-rank pairs in \mathbb{R}^2 . We find *three* settings for the γ 's in which products $\mathbf{x} \cdot \mathbf{y}$ retain full non-degenerate pair status. (Details are in the Appendix.) They represent classes *not* related to each other

by any real coordinate transformation. Taking standard coordinates, these are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \underbrace{\begin{pmatrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix}}_A \quad \text{or} \quad \underbrace{\begin{pmatrix} x_1 y_1 + x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix}}_B \quad \text{or} \quad \underbrace{\begin{pmatrix} x_1 y_1 \\ x_1 y_2 + x_2 y_1 \end{pmatrix}}_C \quad (11)$$

Probing with \mathbf{x} can thus be specified by any of the three 2×2 matrices

$$(\mathbf{x} \cdot) = \underbrace{\begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}}_A \quad \text{or} \quad \underbrace{\begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 \end{pmatrix}}_B \quad \text{or} \quad \underbrace{\begin{pmatrix} x_1 & 0 \\ x_2 & x_1 \end{pmatrix}}_C \quad (12)$$

These varied possibilities allow the richness of physics while limiting the possibilities to those allowed by A and B and C alone.

Lest repeated operation of $(\mathbf{x} \cdot)$ cause targets \mathbf{y} to diverge towards infinity or collapse towards zero, thereby exploding or imploding them, we use $\det(\mathbf{x} \cdot) = 1$, which defines unit quantity of \mathbf{x} . There is no loss of generality because $(\mathbf{x} \cdot)$ can always be rescaled in any particular case. These normalised operators have only one free parameter ϕ related to x_2/x_1 and take the form

$$(\mathbf{x} \cdot) = \underbrace{\begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}}_A \quad \text{or} \quad \underbrace{\begin{pmatrix} \cosh \phi & \sinh \phi \\ \sinh \phi & \cosh \phi \end{pmatrix}}_B \quad \text{or} \quad \underbrace{\begin{pmatrix} 1 & 0 \\ \phi & 1 \end{pmatrix}}_C \quad (13)$$

We can also focus on the nature of the multiplication itself by building ϕ from infinitesimal increments of operators

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{or} \quad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{or} \quad \mathbf{C} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (14)$$

so that

$$(\mathbf{x} \cdot) = \lim_{n \rightarrow \infty} \left(\mathbf{1} + \frac{\phi}{n} \mathbf{G} \right)^n = \exp(\phi \mathbf{G}) \quad (15)$$

where the *generator* \mathbf{G} is the matrix \mathbf{A} or \mathbf{B} or \mathbf{C} .

The first of these (operator A) is rotation by phase ϕ . Taking (7) for addition and (12A) for multiplication, we recognise the sum and product rules of complex arithmetic, so that pairs are complex numbers

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = r e^{i\phi} \quad (16)$$

obeying the standard rules. Moreover, unit quantity is identified with unit determinant, which for (12A) is modulus-squared $\det(\mathbf{x} \cdot) = |\mathbf{x}|^2 = 1$. **Hence the inherent uncertainty in a unit object refers to what remains undefined in \mathbf{x} , namely phase ϕ , so that each new object brings with it an unknown phase.** At this point, *probability* must enter the development.

4 Probability

Our ignorance of phase is uniformly distributed, giving phase a uniform probability distribution

$$\Pr(\phi) = \frac{1}{2\pi}, \quad (0 \leq \phi < 2\pi) \quad (17)$$

around the circle. Otherwise, we could use identically-formed components to build composite objects $\mathbf{X} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n$ whose overall phase would be arbitrarily precise around the supposed mode, thus defeating the engineering requirement of inevitable uncertainty. Meanwhile, phase is necessarily continuous because any restriction would conflict with general application of the complex sum rule. Continuity is not an assumption, it's a requirement.

We have no way of extracting what might be a definitive “truth”. The best we can do in the face of uncertainty is use replication to obtain predictions of

average behaviour, with precision increasing statistically as the square root of the replication factor. These average predictions are called *ensemble averages*.

According to the standard rules of probability, ensemble averaging involves summing (integrating) over all the unknown parameters. Such averaging is known in statistics as “marginalization” by analogy with the treatment of tables of possibilities laid out on a page. In quantum theory here, we average over uniformly distributed phases to get

$$\begin{aligned}\langle |\mathbf{X}|^2 \rangle &= \left\langle |\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n|^2 \right\rangle_{\phi_1, \phi_2, \dots, \phi_n} \\ &= |\mathbf{x}_1|^2 + |\mathbf{x}_2|^2 + \cdots + |\mathbf{x}_n|^2.\end{aligned}\tag{18}$$

Hence by summing over the unknown phases, the quantity we have access to via experiment is the probability, or likelihood, of a given outcome. **This is why quantum mechanics is a probabilistic theory.** This likelihood, which is additive because of the scalar sum rule, is found to refer to the squared modulus, which is the Born rule [23].

It remains to try the alternative operators B and C. Taking (12B) or (12C) for multiplication, we would again identify unit quantity with unit determinant. Uncertainty would again refer to ϕ — now a pseudo-phase — and ignorance of pseudo-phase would again be uniformly distributed.

$$\Pr(\phi) = \text{constant}, \quad (-\infty < \phi < \infty)\tag{19}$$

However, probabilities sum to unity and infinite range makes this normalising constant zero. Such calculus would be unusable because any finite range of ϕ would be infinitely improbable. The engineering would have failed through the distribution (19) being improper, leaving operator A as the only possibility.

This foundation of quantum theory is elementary and simple. Quantity and uncertainty fuse together into complex numbers with uncertainty referring to phase (Fig. 1) and modulus-squared being the observable quantity (the Born rule for arbitrary amount) representing ensemble averages.

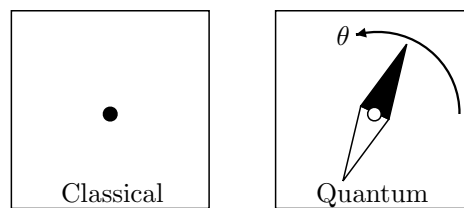


Figure 1: Classical and Quantum objects

5 Inference

Looking ahead and anticipating spacetime which has not yet been developed, physical apparatus produces beams of particles emitted sporadically at some average rate, with intensity determined by averaging over unknown phases. Particles themselves will be secondary to beams and the predictions we make are probabilistic, in the form of an ensemble of possibilities. Standard probability, often redundantly called “Bayesian”, affords a solid foundation for the traditional quantum formalism.

The probabilistic nature of quantum theory has been acknowledged by [24, 25, 26, 27, 28] in “Quantum Bayesianism” (QBism), though it’s actually intrinsic and does not need to be argued or demonstrated separately. Indeed, our formalism was engineered to be automatically consistent with probability theory through obeying the same foundational symmetries [15, 20]. Contradiction

was never possible, so extra assumption was never needed.

Quantum formalism is part of physics: it predicts the likely (probabilistic) behavior of specified models of physical situations. Given those likelihoods, Bayesian analysis then computes posterior probabilities, which assess the models in the light of outcomes as actually observed. But Bayes' Theorem, which formalises our inferences about the world, is not a part of physics. Physics makes predictions quantified probabilistically in terms of likelihoods. The physical world proceeds independently of our thoughts about it (except insofar as we may decide to interrupt its natural evolution). We recommend keeping physics and inference conceptually apart [29].

6 Qubits

So far, we have investigated the calculus of objects whose only relevant property is existence (in some identified state). We now upgrade to objects that can exist in two states, \uparrow and \downarrow say, which we can produce by combining independent \uparrow and \downarrow ensembles. Instead of having a single-state “quantity-with-uncertainty” representation with two real numbers fused into a single complex, we now have a representation involving two complex (four real) numbers.

$$\psi = \begin{pmatrix} \psi_{\uparrow} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi_{\downarrow} \end{pmatrix} = \begin{pmatrix} \psi_{\uparrow} \\ \psi_{\downarrow} \end{pmatrix} = \begin{pmatrix} \psi_0 + i\psi_1 \\ \psi_2 + i\psi_3 \end{pmatrix} \quad (20)$$

Physicists identify such objects as spin-half leptons while mathematicians call them 2-spinors of rank 1. Taking our cue from computation, we call them “qubits” for short. Any number of states greater than two can be decomposed in binary fashion, so using just two loses no generality.

Quantification of \uparrow and \downarrow separately is performed with detectors which are now represented by 2×2 projection matrices

$$\begin{aligned} q_{\uparrow} = |\psi_{\uparrow}|^2 &= \begin{pmatrix} \psi_{\uparrow}^* & \psi_{\downarrow}^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \psi_{\uparrow} \\ \psi_{\downarrow} \end{pmatrix} \\ q_{\downarrow} = |\psi_{\downarrow}|^2 &= \begin{pmatrix} \psi_{\uparrow}^* & \psi_{\downarrow}^* \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \psi_{\uparrow} \\ \psi_{\downarrow} \end{pmatrix} \end{aligned} \quad (21)$$

Under change of basis and linear combination, this generalises to $q = \psi^\dagger \mathbf{M} \psi$ for general quantification of a qubit where \mathbf{M} is an arbitrary matrix — Hermitian by construction and because any anti-Hermitian part would cancel. The observable ensemble average is $\langle q \rangle = \langle \psi^\dagger \mathbf{M} \psi \rangle$, commonly expressed as $\langle q \rangle = \text{trace}(\rho \mathbf{M})$ where $\rho = \langle \psi \psi^\dagger \rangle$ is the *density matrix* though we do not use that construction here.

Under combination, qubits still obey associative distributivity so that their representation is two-dimensional multiplicative as before, except that it's over the complex field instead of the real. We could now be allowed all three of \mathbf{A} and \mathbf{B} and \mathbf{C} , now used as complex-to-complex $\mathbb{C}^2 \times \mathbb{C}^2 \rightarrow \mathbb{C}^2$ operators. In this expanded context there is no reason to exclude \mathbf{B} or \mathbf{C} , but that is the limit of the allowed freedom. Technically, \mathbf{A} and \mathbf{B} become inter-convertible by complex transformation but \mathbf{C} retains its individuality. This implies that \mathbf{C} will provide a quantification distinct from that provided by \mathbf{A} and \mathbf{B} .

Note that ϕ in (13) can also be complex now, so that the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ can have complex coefficients. Leaving \mathbf{C} aside for now, \mathbf{A} and \mathbf{B} yield the Pauli matrices [30, pp. 164-165]

$$\sigma_x = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\mathbf{B}}, \quad \sigma_y = \underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{i\mathbf{A}}, \quad \sigma_z = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\mathbf{BA}} \quad (22)$$

which have been written in Hermitian form for convenience and which (together with the identity matrix $\sigma_0 = \mathbf{1}$) form a closed group under multiplication.

The three Pauli matrices yield quantification

$$\begin{aligned} q_x &= \psi^\dagger \sigma_x \psi = 2(\psi_0 \psi_2 + \psi_1 \psi_3) &= 2 \operatorname{Re}(\psi_\uparrow^* \psi_\downarrow) \\ q_y &= \psi^\dagger \sigma_y \psi = 2(\psi_0 \psi_3 - \psi_1 \psi_2) &= 2 \operatorname{Im}(\psi_\uparrow^* \psi_\downarrow) \\ q_z &= \psi^\dagger \sigma_z \psi = \psi_0^2 + \psi_1^2 - \psi_2^2 - \psi_3^2 &= |\psi_\uparrow|^2 - |\psi_\downarrow|^2 \end{aligned} \quad (23)$$

The overall quantity

$$q_0 = \psi^\dagger \mathbf{1} \psi = \psi_0^2 + \psi_1^2 + \psi_2^2 + \psi_3^2 = |\psi_\uparrow|^2 + |\psi_\downarrow|^2 \quad (24)$$

is not independent, the relationship being

$$q_x^2 + q_y^2 + q_z^2 = q_0^2 \quad (25)$$

which is invariant under ordinary three-dimensional rotation of x, y, z . This is the first clue about the emergence of spacetime from quantum formalism.

On combining independent samples into an ensemble, the vector coordinates q_x, q_y, q_z add and so do their radii q_0 , which allows the equality (25) to degrade to

$$\langle q_x \rangle^2 + \langle q_y \rangle^2 + \langle q_z \rangle^2 \leq \langle q_0 \rangle^2 \quad (26)$$

With complex coefficients ϕ , the three generators (14) define the 6-parameter *Lorentz group* of transformations

$$\psi' = \exp(\phi_x \sigma_x + \phi_y \sigma_y + \phi_z \sigma_z) \psi \quad (27)$$

under which

$$\langle q_0 \rangle^2 - \langle q_x \rangle^2 - \langle q_y \rangle^2 - \langle q_z \rangle^2 \quad \text{is invariant.} \quad (28)$$

7 Spin

Writing the complex coefficients as $\phi = -(\xi + i\eta)/2$, the simplest operation is σ_z with imaginary coefficient $\phi_z = -i\eta/2$, which spins complex ψ_\uparrow and ψ_\downarrow in phase.

$$\begin{pmatrix} \psi'_\uparrow \\ \psi'_\downarrow \end{pmatrix} = \begin{pmatrix} e^{-i\eta/2} & 0 \\ 0 & e^{+i\eta/2} \end{pmatrix} \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} \quad (29)$$

Correspondingly, the observable spins q in (23) and (24) rotate about z as

$$\begin{aligned} q'_0 &= q_0 \\ q'_x &= q_x \cos \eta - q_y \sin \eta \\ q'_y &= q_x \sin \eta + q_y \cos \eta \\ q'_z &= q_z \end{aligned} \quad (30)$$

All spins rotate equally, so the observable ensemble rotates undistorted about z .

In the context of three-dimensional rotation of x, y, z , the ensemble average $\langle q \rangle$ is given symbol $2J$, where J transforms as

$$\begin{pmatrix} J'_0 \\ J'_x \\ J'_y \\ J'_z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \eta & -\sin \eta & 0 \\ 0 & \sin \eta & \cos \eta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} J_0 \\ J_x \\ J_y \\ J_z \end{pmatrix} \quad (31)$$

with the general invariant (28) being $J_0^2 - J_x^2 - J_y^2 - J_z^2$. When, later, space and time have been constructed, J becomes identifiable as angular momentum (including intrinsic spin) in units of \hbar . The observable invariants under rotation are

$$J_0 \quad \text{and} \quad \sqrt{(J_x^2 + J_y^2 + J_z^2)} \quad (32)$$

Although rotation is again recognisable in the equations, this is without explicit reference to space which is yet to be constructed. Observe, though, the

commonly-remarked angle-doubling homomorphism between $\eta/2$ which operates on ψ (group SU(2) rotation of a complex plane) and η which operates on the observable spin q (group SO(3) usually visualised as rotation of a 3-sphere). Rotational invariance under SO(3) has emerged automatically from the formalism, and is not imposed as a property of pre-supposed isotropic space [31, 32].

8 Momentum

We can also apply a real coefficient $\phi_z = -\xi/2$ to σ_z . This rebalances ψ_\uparrow and ψ_\downarrow as

$$\begin{pmatrix} \psi'_\uparrow \\ \psi'_\downarrow \end{pmatrix} = \begin{pmatrix} e^{-\xi/2} & 0 \\ 0 & e^{+\xi/2} \end{pmatrix} \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} \quad (33)$$

Correspondingly, the q 's in (23) and (24) transform as

$$\begin{aligned} q'_0 &= q_0 \cosh \xi - q_z \sinh \xi \\ q'_x &= q_x \\ q'_y &= q_y \\ q'_z &= -q_0 \sinh \xi + q_z \cosh \xi \end{aligned} \quad (34)$$

On replacing the ensemble average $\langle q \rangle$ by symbol p for this new set of real properties, we have

$$\begin{pmatrix} p'_0 \\ p'_x \\ p'_y \\ p'_z \end{pmatrix} = \begin{pmatrix} \cosh \xi & 0 & 0 & -\sinh \xi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh \xi & 0 & 0 & \cosh \xi \end{pmatrix} \begin{pmatrix} p_0 \\ p_x \\ p_y \\ p_z \end{pmatrix} \quad (35)$$

Instead of rotation, this is a Lorentz boost along z , with p_x and p_y unaffected and $p_0^2 - p_z^2$ invariant. Recognising this from relativistic kinematics, we identify p_0 as energy E and (p_x, p_y, p_z) as momentum. Energy and momentum transform as scalar and vector under rotation, and we write the general invariant (28) under boost as

$$E^2 - p_x^2 - p_y^2 - p_z^2 = m^2 \quad (36)$$

where we identify m as (rest-)mass. We see that mass behaves under transformation (35) as a 4-vector with Minkowski metric $\text{diag}(1, -1, -1, -1)$.

Qubits, and consequently all their various combinations, have quantity allied to a directional spin, and energy allied to a directional momentum. Momentum is imaginary spin and spin is imaginary momentum. But we do not yet have the time and space within which energy and momentum are conventionally given meaning.

9 Spacetime

In the complex environment of quantum theory, generators A and B have yielded a real 3-vector called spin and a dual real 3-vector called momentum. Each had an associated scalar making 8 properties in all. Complex 2×2 matrices can support 8 components, but the Hermitian matrices which define observables are only a 4-parameter subset (the Pauli matrices with identity). Something's missing, and it's the third and final allowed multiplication operator C.

Generator C (equation 13C) authorises differentials, thereby allowing integration.

$$\begin{pmatrix} u' \\ U' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \begin{pmatrix} u \\ U \end{pmatrix} = \begin{pmatrix} u \\ U + ut \end{pmatrix} \quad \begin{cases} \delta u = 0 \\ \delta U = u \delta t \end{cases} \quad (37)$$

This generates a new parameter U as the integral of an existing parameter u without changing the latter (Galileo's principle). Conversely, u can be recovered as its differential.

We now inquire into the phase of ψ as in $\psi \propto e^{-i\theta}$ (16), with the minus sign being physicists' convention. For an object of mass m , we can define a scaled

phase τ through

$$d\theta = m d\tau \quad (38)$$

in which we identify θ with U in (37). Phase cannot be rescaled because it's 2π -periodic, so neither can phase differences $\delta\theta$. So, with $d\theta$ being invariant to re-orientation and m transforming as a 4-vector under (36), it follows that $d\tau$ also transforms as a 4-vector under the same Minkowski metric with invariant

$$(dt)^2 - (dx)^2 - (dy)^2 - (dz)^2 = (d\tau)^2 \quad (39)$$

having inner product

$$E dt - p_x dx - p_y dy - p_z dz = d\theta \quad (40)$$

from which we see that phase θ is the physicists' *action* [33]. (In general coordinates, (dt, dx, dy, dz) would be deemed contravariant while (E, p_x, p_y, p_z) would be covariant.)

We now identify τ as proper time, t as clock time, and $\mathbf{x} = (x, y, z)$ as spatial location. **This is why our description of space uses three dimensions.** Velocity defined as $\mathbf{v} = d\mathbf{x}/dt$ is automatically bounded by 1 (the speed of light) through (39), so *Einstein's postulate to that effect is not needed* — it emerges from the formalism. Meanwhile, our descriptions of time and space inherit continuity from necessarily-continuous phase. Continuity is not an assumption, it is a requirement. **Three-dimensional space and one-dimensional time are identified as integrals of Lorentz boosts.** Instead of spacetime being a pre-existing framework into which energy and momentum are inserted, it is a construction emerging from those dynamical observables.

The wave-like nature of matter and energy is due to the periodicity of the phase ϕ (16). With $\psi \propto e^{-i\theta}$, Schrödinger's equation

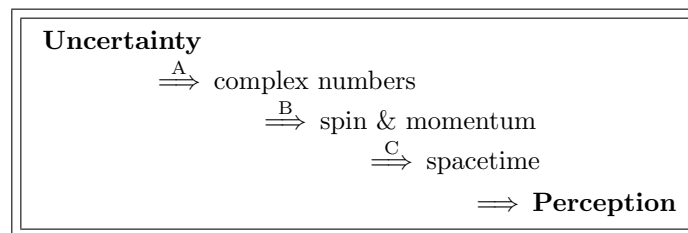
$$i \frac{\partial \psi}{\partial t} = E \psi \quad (41)$$

and the momentum operators

$$-i \frac{\partial \psi}{\partial x} = p_x \psi, \quad -i \frac{\partial \psi}{\partial y} = p_y \psi, \quad -i \frac{\partial \psi}{\partial z} = p_z \psi \quad (42)$$

are immediate, with the linearity of Schrödinger's equation manifest. The factors of i ensure that the operators for energy and momentum are Hermitian, so those properties are observable in accordance with experience.

In summary, relativistic kinematics, with three dimensions of space and one of time, emerges automatically from the foundational structure of quantum theory. This language of physics, in the form of the 6-parameter Lorentz group of spin and momentum, augmented by 4-parameter spacetime, follows inevitably from the necessity of introducing uncertainty at the outset of our inquiry. As macroscopic beings, we are at the end of this chain of development.



Not only do the elementary symmetries of associative commutativity and associative distributivity force the mathematics used to quantify objects and events, but they also force the descriptions underlying our perception of objects and events [34]. These inherent perceptual descriptions lead one to feel that

space and time are an objective reality, whereas our analysis shows that three-dimensional space and one-dimensional time are the only consistent description one can have of reality. Perhaps that's why it has taken centuries for this structure to be uncovered.

It will not escape notice that spatial location is accumulated as the time-integral of local velocity. Consequently, the curvature of space, if any, is not defined in the language, but is part of the content of physics, along with basic parameters of particle physics and cosmology. In this paper, we have addressed the language only, leaving content for other investigation.

10 Conclusions

We have argued that objects must be represented, not by classical scalars, but by pairs of numbers fusing quantity and uncertainty, and have shown how they can be so quantified. Associative commutativity of combination and partition results in the component-wise sum rule, and associative distributivity of chaining or sequencing results in three possible product rules. These define the arithmetical rules to which our formalism must adhere. Scalars have only one product rule but pairs have three.

Only one of the three product rules (A) yields proper (probabilistic) predictions. Applied to pairs, it imposes complex quantum amplitudes controlled by the Feynman sum and product rules with predictions given by the Born rule. By construction, quantum formalism is thus engineered as a probabilistic theory incorporating uncertainty. Uncertainty appears as phase and the theory

is engineered to predict objects' behavior in the form of likelihood functions, which is *physics*. Given those likelihoods, Bayesian analysis then computes posterior probabilities, which assess the models in the light of outcomes as actually observed, which is *inference*. In this way physics and inference are distinct, but mutually consistent, with physics making predictions about the world and inference using those predictions along with experimental outcomes to learn about the world. Of course, there can be no conflict between standard classical probability and quantum applications because the sum and product rules for scalars and pairs follow together from common foundational symmetries to give consistent arithmetical rules.

Every independent component of a compound object has its own unknown phase in a wide-ranging ensemble of possibilities. For large objects, such phases average out leaving classically scalar macroscopic quantities. Where components are not independent because of related construction, their connected phases are said to be “entangled”. Entanglement yields extra information about an object, inaccessible to classical scalar inquiry as exemplified by the Bell inequalities [35, 36].

Beyond phase and entanglement, which are underpinned by product rule A, lie rules B and C which take the scope further. Those rules must also play a part — otherwise some additional assumption would be required to prevent it. Rule B generates the Pauli matrices, which yield spin, energy, and momentum. Rule C integrates those to construct 3+1-dimensional relativistic spacetime. This explains why space has three dimensions, and why Einstein's postulate of

a speed limit (that of light) was correct [37].

Upon acknowledging that quantification carries uncertainty with it, we find that quantum formalism and relativistic spacetime both follow without further assumption. The rules follow from logical derivation, and not from experimentation which must and does conform. As first demonstrated by [11], robust formalism is not invented or discovered, but is openly engineered to conform to fundamental symmetries. There is then no mystery about why the resulting mathematics works. It has to.

We do not, of course, make the philosophically contentious claim to have proved our findings in some absolute sense. All we claim is that any hypothetical alternative theory not conforming to standard quantum formalism and standard spacetime must be incompatible with the foundational symmetries, so that combination and/or sequencing must go awry in that world. In principle, a world could be so deeply interconnected that separating individual objects was impossible. In practice, ours isn't, so we stand by our findings as a uniquely consistent mathematical description of the world we live in.

Scalars:

Shuffling := commutative + associative		
Sequencing := distributive + associative		
Shuffling	→	Sum rule, quantification } Measures
Shuffling	→	Sum rule, quantity
Sequencing	→	Product rule, proportion } Probability

Pairs:

Quantity & uncertainty → represent by number pair	
Shuffling	→ component-wise sum rule
Sequencing	→ three allowed product rules A,B,C A = complex multiplication B = hyperbolic variant C = integrator
1 property (existence) → $\left\{ \begin{array}{l} \text{Only A is consistent with probability} \\ \therefore \text{ pair = } \mathbf{complex\ number\ } \psi, \text{ with} \\ \text{complex + and } \times \text{ (} \mathbf{Feynman\ picture} \text{)} \\ \text{and quantity = } \psi ^2 \text{ (the } \mathbf{Born\ rule} \text{)} \end{array} \right.$	
2 properties (qubit) → $\left\{ \begin{array}{l} \text{A,B,C can all apply} \\ \text{A,B} \rightarrow \text{Pauli matrices} \rightarrow \mathbf{spin, 4-momentum} \\ \text{C on 4-momentum} \rightarrow \mathbf{relativistic\ spacetime} \end{array} \right.$	

Acknowledgments

The authors thank Hans Eggers, Ariel Caticha, Seth Chaiken, Keith Earle, Cecilia Levy, Oleg Lunin, Matthew Szydagis, Steve Gull, Anton Garrett, and Chris Doran — also Dylan VanAllen and the Spring 2020 Advanced Physics Laboratory (APHY 335Z) and Fall 2020 Bayesian Data Analysis (APHY/ICSI/IINF 451/551) classes at the University at Albany — for their helpful discussions on the material related to and presented in this paper. Knuth also thanks the Foundational Questions Institute (FQXi) for their Essay Contests, which have encouraged many of us physicists to carefully think and write about these deep issues.

References

- [1] E. P. Wigner, *Communications on pure and applied mathematics* **13**, 1 (1960). Richard Courant lecture in mathematical sciences delivered at New York University, May 11, 1959.
- [2] R. W. Hamming, *American Mathematical Monthly* pp. 81–90 (1980).
- [3] R. Feynman, *The character of physical law (1965)* (MIT Press, 1967).
- [4] A. I. Miller, *Sixty-Two Years of Uncertainty: historical, philosophical, and physical inquiries into the foundations of quantum mechanics*, vol. 226 (Springer Science & Business Media, 2012).
- [5] C. G. Timpson, *Quantum information theory and the foundations of quantum mechanics* (OUP Oxford, 2013).
- [6] L. Hardy, *Quantum theory: informational foundations and foils* (Springer, 2016), pp. 223–248.
- [7] R. Friedberg, P. C. Hohenberg, *Foundations of Physics* **48**, 295 (2018).
- [8] A. Caticha, *Entropy* **21**, 943 (2019).
- [9] G. Jaeger, *Annalen der Physik* **531**, 1800097 (2019).
- [10] J. Pfanzagl, *Theory of measurement* (Wiley, 1968).
- [11] R. T. Cox, *Am. J. Physics* **14**, 1 (1946).
- [12] K. H. Knuth, *Trick or Truth: the Mysterious Connection between Physics and Mathematics*, A. Aguirre, B. Foster, Z. Merali, eds., Springer Frontiers

Collection (Springer-Verlag, Heidelberg, 2016), pp. 77–90. FQXi 2015 Essay Contest (Third Prize),(arXiv:1504.06686 [math.HO]).

- [13] J. Aczél, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Jackson Hole WY, USA, August 2003, G. J. Erickson, Y. Zhai, eds., AIP Conf. Proc. 707 (AIP, New York, 2004), pp. 195–203.
- [14] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge Univ. Press, Cambridge, 2003).
- [15] K. H. Knuth, J. Skilling, *Axioms* **1**, 38 (2012).
- [16] C. R. Smith, G. J. Erickson, *Maximum Entropy and Bayesian Methods*, Dartmouth USA, 1989, P. F. Fougère, ed. (Kluwer, Dordrecht, 1990), pp. 17–30.
- [17] A. Caticha, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Oxford MS, USA 2009, P. M. Goggans, C.-Y. Chan, eds. (AIP, 2009), vol. 1193, pp. 60–68.
- [18] K. H. Knuth, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Jackson Hole WY, USA, August 2003, G. J. Erickson, Y. Zhai, eds., AIP Conf. Proc. 707 (AIP, New York, 2004), pp. 204–235. (arXiv:physics/0403031v1 [physics.data-an]).
- [19] K. H. Knuth, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Oxford, MS, USA, 2009, P. Goggans, C.-Y.

- Chan, eds., AIP Conf. Proc. 1193 (AIP, New York, 2009), pp. 132–144.
(arXiv:0909.3684v1 [math.GM]).
- [20] J. Skilling, K. H. Knuth, *Annalen der Physik* **531**, 1800057 (2019).
- [21] B. de Finetti, *Biblioteca di Filosofia* (Libreria editrice Perrella, 1931).
- [22] C. M. Caves, (*unpublished*) <http://info.phys.unm.edu/~caves/reports/dutchbook.pdf> (2000).
- [23] M. Born, *Zeit fur Phys* **38**, 803 (1926).
- [24] C. M. Caves, C. A. Fuchs, R. Schack, *Phys. Rev. A* **65**, 022305 (2002).
- [25] C. A. Fuchs, Quantum mechanics as quantum information (and only a little more) (2002). (arXiv:quant-ph/0205039).
- [26] I. Pitowsky, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **34**, 395 (2003). Quantum Information and Computation.
- [27] J. Bub, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **38**, 232 (2007).
- [28] C. A. Fuchs, QBism, the perimeter of quantum Bayesianism (2010). (arXiv:1003.5209 [quant-ph]).
- [29] E. T. Jaynes, *Maximum Entropy and Bayesian Methods in Science and Engineering* (Springer, 1989), pp. 1–27.

- [30] J. J. Sakurai, *Modern quantum mechanics, revised edition*, S. F. Tuan, ed. (Addison-Wesley Publishing Co, 1994).
- [31] R. Penrose, W. Rindler, *Spinors and space-time: Volume 1, Two-spinor calculus and relativistic fields*, vol. 1 (Cambridge University Press, 1984).
- [32] S. Gull, A. Lasenby, C. Doran, *Foundations of Physics* **23**, 1175 (1993).
- [33] R. P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).
- [34] A. Pellionisz, R. Llinás, *Neuroscience* **7**, 2949 (1982).
- [35] J. S. Bell, *Physics* **1**, 195 (1964).
- [36] R. F. Werner, M. M. Wolf, *Quantum Info. Comput.* **1**, 1 (2001).
- [37] A. Einstein, *Annalen der Physik.* **17**, 891 (1905).
- [38] P. Goyal, K. H. Knuth, J. Skilling, *Phys. Rev. A* **81**, 022109 (2010).
(arXiv:0907.0909 [quant-ph]).

Appendix

This analysis updates and clarifies material originally presented in Appendix A of the GKS paper [38]. In that paper, the primes in its equation (19) were inconsequentially misplaced — they should have been on its right-hand γ 's, not the left.

Associative commutativity of $+$ requires combination $\mathbf{c} = \mathbf{a} + \mathbf{b}$ of n -tuples to be represented — up to isomorphism — by component-wise addition (the

sum rule, which is unique for any n).

$$\left. \begin{array}{l} \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \quad (\text{commutative}) \\ (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}) \quad (\text{associative}) \end{array} \right\} \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \\ \implies c_i = a_i + b_i \quad \forall i$$

This linear form is invariant to non-singular linear transformation of axes. Distributivity of \cdot over $+$ then requires n -tuple connection $\mathbf{c} = \mathbf{a} \cdot \mathbf{b}$ to be bilinear multiplication defined by n^3 coefficients γ .

$$\left. \begin{array}{l} \mathbf{z} \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{z} \cdot \mathbf{x} + \mathbf{z} \cdot \mathbf{y} \quad (\text{left distributive}) \\ (\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z} \quad (\text{right distributive}) \end{array} \right\} \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \\ \implies c_i = \sum_{j=1}^n \sum_{k=1}^n \gamma_{ijk} a_j b_k \quad \forall i$$

Associative distributivity imposes n^4 quadratic constraints (not all independent) on those n^3 γ 's.

$$(\mathbf{x} \cdot \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot (\mathbf{y} \cdot \mathbf{z}) \quad (\text{associative distributivity}) \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \\ \implies \sum_{t=1}^n (\gamma_{tpq} \gamma_{rst} - \gamma_{tsp} \gamma_{rtq}) = 0 \quad \forall p, q, r, s$$

The associativity is to be nontrivial: $\mathbf{x} \cdot \mathbf{y} \cdot \mathbf{z} \neq \mathbf{0}$. Associative commutativity and associative distributivity are our foundational symmetries. The constraints are simple quadratics, which accords with intuition that generality should be simple.

We seek standardised values for the multiplication coefficients γ that we can adopt as agreed conventions for representing families connected by non-singular transformation of axes. Each such standard will be called a *product rule*.

For a start, we can scale axes by

$$x_j \longrightarrow x'_j = x_j / \lambda \quad \implies \quad \gamma'_{jjj} = \lambda \gamma_{jjj}$$

for each index j separately. This lets us set the coefficients γ_{jjj} (all 3 suffices the same) to 0 or +1 for each j .

One dimension $n = 1$

With only one coefficient γ_{111} , setting it to 0 would be trivial and useless so we assign

$$\gamma_{111} = 1$$

Alternative conventions are possible, as when proportions are presented as percentages ($\gamma_{111} = 1/100$), but the unit assignment is standard. In one dimension, distributivity allows *only one product rule* (which is ordinary scalar multiplication) and associativity follows trivially. But in higher dimension more than one product rule can accompany the unique sum rule.

Two dimensions $n = 2$

There are now $2^3 = 8$ γ 's, which we lay out as an array

$$\gamma = \begin{bmatrix} \gamma_{111} & \gamma_{112} & \gamma_{121} & \gamma_{122} \\ \gamma_{211} & \gamma_{212} & \gamma_{221} & \gamma_{222} \end{bmatrix}$$

The $2^4 = 16$ associativity constraints are

$$\gamma_{111}\gamma_{111} - \gamma_{111}\gamma_{111} + \gamma_{211}\gamma_{112} - \gamma_{211}\gamma_{121} = 0 \quad (1)$$

$$\gamma_{111}\gamma_{121} - \gamma_{121}\gamma_{111} + \gamma_{211}\gamma_{122} - \gamma_{221}\gamma_{121} = 0 \quad (2)$$

$$\gamma_{111}\gamma_{211} - \gamma_{111}\gamma_{211} + \gamma_{211}\gamma_{212} - \gamma_{211}\gamma_{221} = 0 \quad (3)$$

$$\gamma_{111}\gamma_{221} - \gamma_{121}\gamma_{211} + \gamma_{211}\gamma_{222} - \gamma_{221}\gamma_{221} = 0 \quad (4)$$

$$\gamma_{112}\gamma_{111} - \gamma_{111}\gamma_{112} + \gamma_{212}\gamma_{112} - \gamma_{211}\gamma_{122} = 0 \quad (5)$$

$$\gamma_{112}\gamma_{121} - \gamma_{121}\gamma_{112} + \gamma_{212}\gamma_{122} - \gamma_{221}\gamma_{122} = 0 \quad (6)$$

$$\gamma_{112}\gamma_{211} - \gamma_{111}\gamma_{212} + \gamma_{212}\gamma_{212} - \gamma_{211}\gamma_{222} = 0 \quad (7)$$

$$\gamma_{112}\gamma_{221} - \gamma_{121}\gamma_{212} + \gamma_{212}\gamma_{222} - \gamma_{221}\gamma_{222} = 0 \quad (8)$$

$$\gamma_{121}\gamma_{111} - \gamma_{112}\gamma_{111} + \gamma_{221}\gamma_{112} - \gamma_{212}\gamma_{121} = 0 \quad (9)$$

$$\gamma_{121}\gamma_{121} - \gamma_{122}\gamma_{111} + \gamma_{221}\gamma_{122} - \gamma_{222}\gamma_{121} = 0 \quad (10)$$

$$\gamma_{121}\gamma_{211} - \gamma_{112}\gamma_{211} + \gamma_{221}\gamma_{212} - \gamma_{212}\gamma_{221} = 0 \quad (11)$$

$$\gamma_{121}\gamma_{221} - \gamma_{122}\gamma_{211} + \gamma_{221}\gamma_{222} - \gamma_{222}\gamma_{221} = 0 \quad (12)$$

$$\gamma_{122}\gamma_{111} - \gamma_{112}\gamma_{112} + \gamma_{222}\gamma_{112} - \gamma_{212}\gamma_{122} = 0 \quad (13)$$

$$\gamma_{122}\gamma_{121} - \gamma_{122}\gamma_{112} + \gamma_{222}\gamma_{122} - \gamma_{222}\gamma_{122} = 0 \quad (14)$$

$$\gamma_{122}\gamma_{211} - \gamma_{112}\gamma_{212} + \gamma_{222}\gamma_{212} - \gamma_{212}\gamma_{222} = 0 \quad (15)$$

$$\gamma_{122}\gamma_{221} - \gamma_{122}\gamma_{212} + \gamma_{222}\gamma_{222} - \gamma_{222}\gamma_{222} = 0 \quad (16)$$

As well as scaling, we can also shear one of the axes by the other, as in

$$x_2 \longrightarrow x'_2 = x_2 - \alpha x_1.$$

The γ 's then transform with, in particular,

$$\begin{aligned}\gamma'_{211} &= \gamma_{211} + (\gamma_{221} + \gamma_{212} - \gamma_{111})\alpha \\ &\quad + (\gamma_{222} - \gamma_{121} - \gamma_{112})\alpha^2 - \gamma_{122}\alpha^3\end{aligned}$$

Full-rank cubic equations have at least one real root, enabling us to set $\gamma_{211} = 0$ unless $\gamma_{122} = 0$. Finally, we can interchange axis labels $1 \rightleftharpoons 2$, so consequently we can always select $\underline{\gamma_{211} = 0}$.

We now choose γ_{111} to be either 1 or 0.

Two dimensions $n = 2$, Choice #1.

On appending $\underline{\gamma_{111} = 1}$ to the original $\underline{\gamma_{211} = 0}$, the associativity equations reduce to

$$\begin{aligned}\gamma_{121}\gamma_{221} &= 0 & (2') \\ \gamma_{112}\gamma_{212} &= 0 & (5') \\ \gamma_{221}(\gamma_{221} - 1) &= 0 & (4') \\ \gamma_{212}(\gamma_{212} - 1) &= 0 & (7') \\ \gamma_{122}(\gamma_{212} - \gamma_{221}) &= 0 & (6') \\ \gamma_{122}(\gamma_{121} - \gamma_{112}) &= 0 & (14') \\ \gamma_{112}(\gamma_{221} - 1) &= \gamma_{121}(\gamma_{212} - 1) & (9') \\ \gamma_{122}(\gamma_{221} - 1) &= \gamma_{121}(\gamma_{222} - \gamma_{121}) & (10') \\ \gamma_{122}(\gamma_{212} - 1) &= \gamma_{112}(\gamma_{222} - \gamma_{112}) & (13') \\ \gamma_{212}(\gamma_{222} - \gamma_{121}) &= \gamma_{221}(\gamma_{222} - \gamma_{112}) & (8')\end{aligned}$$

From (4'), $\gamma_{221} = 1$ or 0; and from (7'), $\gamma_{212} = 1$ or 0. This allows four choices (#1.1, #1.2, #1.3, #1.4).

Two dimensions $n = 2$, Choice #1.1.

Choice #1.1 adds $\underline{\gamma_{221} = 1}$ and $\underline{\gamma_{212} = 1}$ to $\gamma_{111} = 1$ and $\gamma_{211} = 0$.

From (2'), $\gamma_{121} = 0$; and from (5'), $\gamma_{112} = 0$. All equations are then satisfied, with γ taking the form

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & \theta \\ 0 & 1 & 1 & \phi \end{bmatrix}$$

Free parameter ϕ can be sheared away by $x'_1 = x_1 + \frac{1}{2}\phi x_2$, giving

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & \theta' \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

On rescaling the “2” axis, $\gamma'_{122} = \theta'$ can be set to standard values of -1 (if originally negative) or 1 (if originally positive) or 0 (if originally zero). This gives three standard forms, similar in style but not related by any real coordinate transformation:

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad (\text{A,B,C})$$

Two dimensions $n = 2$, Choice #1.2.

Choice #1.2 adds $\gamma_{221} = 0$ and $\gamma_{212} = 1$ to $\gamma_{111} = 1$ and $\gamma_{211} = 0$.

From (5'), $\gamma_{112} = 0$; from (6'), $\gamma_{122} = 0$; and from (8'), $\gamma_{121} = \gamma_{222}$. All equations are then satisfied, with γ taking the form

$$\gamma = \begin{bmatrix} 1 & 0 & \phi & 0 \\ 0 & 1 & 0 & \phi \end{bmatrix}$$

Free parameter ϕ can be sheared away by $x'_1 = x_1 + \phi x_2$ to give the standard form

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (\text{D})$$

Two dimensions $n = 2$, Choice #1.3.

Choice #1.3 adds $\gamma_{221} = 1$ and $\gamma_{212} = 0$ to $\gamma_{111} = 1$ and $\gamma_{211} = 0$.

This is the $\mathbf{a} \cdot \mathbf{b} \Rightarrow \mathbf{b} \cdot \mathbf{a}$ reversal $\gamma_{ijk} \Rightarrow \gamma_{ikj}$ of choice #1.2 and leads to the reversed standard form

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (\text{E})$$

Two dimensions $n = 2$, Choice #1.4.

Choice #1.4 adds $\gamma_{221} = 0$ and $\gamma_{212} = 0$ to $\gamma_{111} = 1$ and $\gamma_{211} = 0$.

From (9'), $\gamma_{121} = \gamma_{112}$; and from (10'), $\gamma_{122} = \gamma_{121}(\gamma_{121} - \gamma_{222})$. All equations are then satisfied, with γ taking the form

$$\gamma = \begin{bmatrix} 1 & \theta & \theta & \theta(\theta - \phi) \\ 0 & 0 & 0 & \phi \end{bmatrix}$$

Free parameter θ can be sheared away by $x'_1 = x_1 + \theta x_2$ to give

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi \end{bmatrix}$$

Scaling $\gamma_{222} = \phi$ to 0 or 1 then appears to give two standard forms

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{F, G})$$

However, form G shears under $(x'_1, x'_2) = \frac{1}{2}(x_1 - x_2, x_1 + x_2)$ to form B above, so is not new.

Two dimensions $n = 2$, Choice #2.

On adding $\underline{\gamma_{111} = 0}$ to the original $\underline{\gamma_{211} = 0}$, the associativity equations reduce

to

$$\begin{aligned}
 \gamma_{121}\gamma_{221} &= 0 & (2'') \\
 \gamma_{112}\gamma_{212} &= 0 & (5'') \\
 \gamma_{221}\gamma_{221} &= 0 & (4'') \\
 \gamma_{212}\gamma_{212} &= 0 & (7'') \\
 \gamma_{112}\gamma_{221} &= \gamma_{121}\gamma_{212} & (9'') \\
 \gamma_{122}(\gamma_{212} - \gamma_{221}) &= 0 & (6'') \\
 \gamma_{122}(\gamma_{121} - \gamma_{112}) &= 0 & (14'') \\
 \gamma_{121}(\gamma_{222} - \gamma_{121}) &= \gamma_{122}\gamma_{221} & (10'') \\
 \gamma_{112}(\gamma_{222} - \gamma_{112}) &= \gamma_{122}\gamma_{212} & (13'') \\
 \gamma_{221}(\gamma_{112} - \gamma_{222}) &= \gamma_{212}(\gamma_{121} - \gamma_{222}) & (8'')
 \end{aligned}$$

From (4''), $\gamma_{221} = 0$; and from (7''), $\gamma_{212} = 0$; leaving only

$$\begin{aligned}
 \gamma_{122}(\gamma_{121} - \gamma_{112}) &= 0 & (14''') \\
 \gamma_{121}(\gamma_{121} - \gamma_{222}) &= 0 & (10''') \\
 \gamma_{112}(\gamma_{112} - \gamma_{222}) &= 0 & (13''')
 \end{aligned}$$

From (10''), γ_{121} is 0 or γ_{222} ; and from (13''), γ_{112} is 0 or γ_{222} ; allowing four forms

$$\begin{aligned}
 \gamma = \begin{bmatrix} 0 & \theta & \theta & \phi \\ 0 & 0 & 0 & \theta \end{bmatrix} & \text{ and } \begin{bmatrix} 0 & 0 & 0 & \phi \\ 0 & 0 & 0 & \theta \end{bmatrix} \\
 & \text{ and } \begin{bmatrix} 0 & \theta & 0 & \phi \\ 0 & 0 & 0 & \theta \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & \theta & \phi \\ 0 & 0 & 0 & \theta \end{bmatrix}
 \end{aligned}$$

To be nontrivial, $\mathbf{x} \cdot \mathbf{y} \cdot \mathbf{z} \not\equiv \mathbf{0}$, these forms need $\gamma_{222} = \theta \neq 0$. Hence θ must be scalable to +1. Equation (14''') then requires $\phi = 0$ in the last two forms, which reproduce earlier solutions D and E in interchanged $1 \rightleftharpoons 2$ form, so offer nothing new. Meanwhile ϕ can be sheared to 0 in the first two forms by $x'_1 = x_1 \pm \phi x_2$ (with + sign for the first, - for the second). Again, these reproduce earlier solutions C and F in interchanged form, so offer nothing new.

Summary

In two dimensions, associative distributivity allows *six standard product rules*, all different, which we label alphabetically.

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 - a_2 b_2 \\ a_1 b_2 + a_2 b_1 \end{pmatrix} \quad (\text{A})$$

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 + a_2 b_2 \\ a_1 b_2 + a_2 b_1 \end{pmatrix} \quad (\text{B})$$

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 + a_2 b_1 \end{pmatrix} \quad (\text{C})$$

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (\text{D})$$

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} b_1 \quad (\text{E})$$

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ 0 \end{pmatrix} \quad (\text{F})$$

Only the first three offer the nondegenerate “Pair · Pair → Pair” product laws that we seek to engineer. Rules D and E introduce multiplication by scalars, and F confirms ordinary multiplication of those scalars. Thus the full structure of an associative algebra follows from pair-wise associative commutativity and nontrivial associative distributivity alone, with scalars emerging automatically as pairs of form $(x, 0)$.