The Born Exponent from Counting: Quantum Probability in a Finite Configuration Ontology

Russell Tillitt
Independent Researcher, San Francisco, California

August 2026

Abstract

We prove that, in a finite and timeless configuration ontology with relational phase structure, the population of an outcome class—the number of configurations compatible with an observer's records—equals $N|\alpha|^2$, where N is the population of the parent class and α is the coherent sum of stage-relative configuration phases over the outcome class. Because probability in this ontology is *defined* as a counting ratio, the Born rule follows: quantum probability is quadratic because counting is quadratic. The mathematical core of the argument is classical; the contribution is the premise: the conserved quantity is a cardinality, not a normalization convention. Phase enters the ontology at two levels: it attaches additively to degrees of freedom, making phase additivity a lemma rather than an assumption; and it is relational, defined for each configuration relative to a stage of an observer chain—we prove that no context-free phase assignment is consistent with the existence of nontrivial devices. Given the linear composition of devices, the derivation proceeds in two steps, each by contradiction. First, a population rule that is not a pure power of $|\alpha|$ contradicts the multiplicativity of counting for independent subsystems. Second, an exponent other than two contradicts conservation of the number of configurations, which holds automatically because a device stands between two stages whose records differ by no outcome information and therefore share one fixed compatibility class. A necessity argument shows that an ontic Born exponent requires canonically measurable—hence finite—outcome classes; passing from populations to observed frequencies requires exactly one explicitly stated typicality assumption, on which the value of the exponent does not depend. All results are verified by literal enumeration in an explicit finite model with class populations up to 4.2×10^6, and finiteness yields falsifiable departures from continuum quantum mechanics: every probability is rational, and Born statistics hold to $O(1/N)$.

1 Introduction

Quantum mechanics assigns probabilities by the Born rule: the probability of an outcome is the squared magnitude of its amplitude. In standard formulations the rule is postulated. Existing derivations supply epistemic or operational principles to fix the quadratic form: rationality constraints on agents [9, 27, 28], entanglement-assisted symmetries of credences [31], indifference over self-location [24], or noncontextuality on a presupposed Hilbert-space structure [14, 6]. In the Everettian setting, the direct alternative—counting branches—fails [15], and the diagnosis is structural: branch number is coarse-graining-relative and undefined [28]. There is no denominator.

This paper works in an ontology that restores the denominator: a finite, timeless configuration space, in which probability is *defined* as counting and the class sizes that counting requires are canonical cardinalities. Within that ontology, we prove that the counting is quadratic in the coherent phase sum.

The construction is a principle theory in Einstein's sense [12]: empirically established regularities are elevated to postulates, and the derivation exhibits what any consistent physics must look like given them—the method of thermodynamics and of 1905 relativity. The empirical principle here is that interference occurs, formalized as the nontriviality clause of [LC] (Section 3) and standing to this paper as the light postulate stands to special relativity; the ontological commitment is finiteness; and the named idealizations—invertibility, attainability—are stated as such and realized to $O(1/N)$ in the explicit model. The derivation's shape is therefore: measured law plus named idealizations yields a unique exponent. In any such ontology, either no interference exists at all, or populations obey $n_i = N|\alpha_i|^2$.

What is proved, and what is not. Two claims must be kept distinct, and this paper keeps them distinct throughout.

- *The Born exponent* (Theorems 1–3): the population of an outcome class—a cardinality, a fact about the world—equals $N|\alpha|^2$. This is proved from the postulates and the structural assumptions of Section 3, with no principle about credence, rationality, or typicality anywhere in the argument.

- *The Born rule as an empirical frequency law* (Corollary 2): that the outcome frequencies an observer actually records track the population fractions. This consumes exactly one further principle, stated up front as Assumption [T] (Section 2.2): typical record contents track the population ratios. [T] is the counting analog of Bohmian quantum equilibrium [11], in the weakest form available—typicality by literal cardinality over equally real fundamental entities—and it is an assumption, not a theorem, here as everywhere.

The value of the exponent does not depend on [T]. The title of this paper names the first claim because the first claim is what is proved outright.

Scope. The result is conditional: *given* the postulates of Section 2 and the assumptions of Section 3, the quadratic law follows. The assumptions are not derived here; in particular [LC], the linear composition of devices, is where the linear structure of quantum mechanics enters this paper (Section 3.1), and its origin belongs to the reconstruction program [17, 8, 21]. The demonstrated regime is finite-dimensional, discrete-phase, interferometric quantum mechanics: preparations, beamsplitter-type devices, detection. Continuous observables and Hilbert-space dynamics belong to a continuum-limit program developed in the companion framework paper [26]. The Born relation is proved for outcome partitions that can support stable records (Theorem 4); partitions that cannot carry no observer-relative outcome facts, and the rule makes no claim about them. In a literally finite ontology the attainability idealization of [LC] holds up to $O(1/N)$ granularity, and the theorems hold correspondingly in robust form (Remark 5). Further limits, including the open problem of base-level phase texture, are collected in Section 11.

Plan. Section 2 states the ontology—including its two-level phase structure—and defines probability as counting, closing with the status of the probability–population identification and Assumption [T]. Section 3 collects the structural assumptions and locates where the quantum structure enters. Section 4 derives the form of the population rule and Section 5 the exponent; Section 6 assembles the Born theorem and its interference corollaries; Section 7 fixes the domain; Section 8 reports the computational verification; Section 9 tabulates the status of every claim; Section 10 relates the result to prior work; Section 11 states limitations and concludes.

2 Ontology and Definitions

Postulate 1 (Finite timeless configuration space with relational phase)**.** Physical reality consists of a timeless configuration space Ω with $|\Omega| < \infty$. Each configuration $\omega \in \Omega$ specifies a complete arrangement of fundamental degrees of freedom. Configurations do not evolve, and none is created or destroyed.

Configuration space carries phase structure, at two levels.

- *Phase attaches to degrees of freedom.* The arrangements of individual degrees of freedom carry phase, and the phase content of a configuration is composed additively from the phase contents of its independent parts.
- *Phase is relational.* No configuration carries an absolute, context-free phase. What exists is the phase of a configuration *relative to a stage of an observer chain*: an angle $S_r(\omega)$, defined modulo 2π for $\omega \in \Omega_r$, accumulated along the closure structure connecting the stage's correlational content to ω. Only phase differences within a stage are physical, and successive stages of a chain are related by increments: $S_{r'}(\omega) = S_r(\omega) + \Delta S_{r\to r'}(\omega)$.

Remark 1. Three consequences of the phase postulate are drawn below rather than assumed: phase additivity for independent subsystems is Lemma 1; the impossibility of a context-free assignment is Remark 4, so the relational clause is forced, not stylistic; and the packaging of phases as unimodular exponentials $e^{iS_r(\omega)}$, summed coherently, is derived in Lemma 3.

Postulate 2 (Records and compatibility classes)**.** Information exists only as records: finite arrangements of degrees of freedom whose structure corresponds to correlations. A record r defines its compatibility class

$$\Omega_r = \{\, \omega \in \Omega : \omega \text{ is compatible with the correlations encoded in } r \,\}, \tag{1}$$

whose members are, by definition, indistinguishable with respect to r.

Definition 1 (Observers and observer chains)**.** An observer stage is a structured subsystem within a configuration whose degrees of freedom include records. Observers are defined structurally; no assumption of consciousness, experience, or agency is made, and "observer" may be read throughout as "record-bearing subsystem." An ordered family of structurally related observer stages, borne by distinct configurations and ordered by refinement of their records, is an observer chain. "The observer" denotes the identification across a chain: a family of subsystems of distinct configurations connected by closure structure, the records of each stage refining and representing the correlational content of prior stages. This definition is vocabulary only; that the world answers to it is the next postulate.

Postulate 3 (Existence of observer chains)**.** Configurations exist that contain structurally related observer stages whose records are ordered by refinement: observer chains exist.

Remark 2. The existence clause is stated as a postulate because it can fail: a model of Postulates 1–2 could contain no record-bearing subsystems, or none ordered by refinement, and Definition 2 below would then be vacuously satisfied. Folding the clause inside the definition of observers would leave it unexamined: a definition cannot be false, and this can. The explicit model of Section 8 realizes chains by construction.

Remark 3 (Sequential language). Nothing in this ontology moves or evolves. Change is difference read along a refinement ordering; time is the emergent coordinate labeling that ordering (developed in the companion paper). All sequential language in this paper—"a device is applied," "a record is inserted," "downstream"—is shorthand for position in the refinement ordering of an observer chain, never for temporal process. The shorthand is used throughout.

2.1 Phase structure

Lemma 1 (Phase additivity). *For independent degrees of freedom, phases add: $S_r(\omega_A, \omega_B) = S_r(\omega_A) + S_r(\omega_B)$ at every stage r whose records leave the two uncorrelated. Two subsystems are independent when no correlation constraint links their degrees of freedom, so that the joint configuration space is the Cartesian product.*

Proof. Immediate from the first clause of Postulate 1: a configuration's phase content composes additively from the phase contents of its independent parts, and stage-relative accumulation respects the composition because closure increments act on correlational content, which for uncorrelated subsystems factorizes. □

Remark 4 (No context-free phase assignment is possible). The relational clause of Postulate 1 is forced by the rest of the framework, in the following precise sense. Suppose instead that each configuration carried a single context-free phase $S(\omega)$, feeding the coherent sums of *every* partition of a class. Any two partitions of one class then have equal totals: summing $e^{iS(\omega)}$ over the whole class gives the same number whether grouped by one partition or the other. Now take a balanced interferometer in this ontology (Section 5): the upstream partition (arms, with arrows proportional to 1 and $e^{i\delta}$) and the downstream partition (ports, with arrows proportional to $1 + e^{i\delta}$ and $1 - e^{i\delta}$, up to the common normalization) are partitions of one class—no outcome information separates the stages. The port total is proportional to 2; the arm total to $\sqrt{2}\,(1+e^{i\delta})$; these are unequal for generic δ. A context-free assignment therefore admits no device whose arithmetic is the interferometer's—which is to say, no nontrivial device at all. Relational, stage-indexed phase is the only reading of "only phase differences are physical" under which nontrivial devices exist: a difference needs a relatum, and the relatum is position in the refinement ordering. (The explicit finite model of Section 8 carries its phases stage-relative; this remark is why it must.)

Definition 2 (Probability as counting). Let r be the record of an observer stage, and let $\{r_i\}_{i=1}^k$ be a family of refinements of r borne by later stages of observer chains extending that stage, with $\Omega_{r_i} \subseteq \Omega_r$. The probability of outcome i, relative to the stage bearing r, is

$$P(r_i \mid r) = \frac{|\Omega_{r_i}|}{|\Omega_r|}. \tag{2}$$

A measurement is an ordered pair of stages whose records are related by refinement and whose compatibility classes differ; measurement update is elimination: the later stage's class is a subset of the earlier stage's, and conditional probabilities are counts over the configurations that remain.

Throughout this paper, N denotes $|\Omega_r|$, the population of the class being partitioned; the total $|\Omega|$ plays no role in the derivation (a fact whose significance is drawn in Proposition 1).

Definition 3 (Outcome partitions and populations)**.** A preparation is a correlation structure: a constraint on which configurations exist and what phases they carry relative to the preparing stage. Parameters of laboratory devices enter only through this structure, never as weights on configurations. An outcome partition of Ω_r is a partition into refinements $\{r_i\}$, mutually exclusive and jointly exhaustive. We call $n_i = |\Omega_{r_i}|$ the *population* of class r_i: the number of configurations it contains. Populations are cardinalities, and every probabilistic statement in this paper is a ratio of populations.

2.2 The identification of probability with population, and Assumption [T]

Definition 2 identifies probability with population fraction; because the identification carries philosophical weight, its status is stated up front.

The problem every derivation of the Born rule must solve is that its ontology contains no probability-shaped quantity: branch structures, insofar as they are countable at all, count evenly, and unequal probabilities must then be supplied from outside by principles about credence. The present ontology contains the graded quantity: population. It is objective, canonical (a cardinality), additive over exclusive alternatives, normalized ($\sum_i n_i = N$), multiplicative over independent subsystems, and—this is the theorem of this paper—equal to $N|\alpha_i|^2$. Identifying probability with population fraction is therefore a functional-role identification, of the same character as identifying temperature with mean kinetic energy: the quantity occupies the role completely.

One component of the role is not fixed by the ontology, and it is stated here as a named assumption:

Assumption 1 ([T] Typicality)**.** Typical record contents track the population ratios: among all configurations containing stored k-run outcome counts, the population is overwhelmingly concentrated on counts near n_i/N (a counting statement, established under the conditions of Corollary 2), and the sub-populations bearing deviant stored counts are not the ones to reason from (the principle).

The first clause of [T] is provable by counting where its conditions hold; the second is not a theorem, here or anywhere. [T] is the counting analog of quantum equilibrium in Bohmian mechanics [11], in the weakest form available: typicality by literal cardinality over equally real fundamental entities, with no measure to choose. Configurations bearing deviant records are a *prediction* of the framework, not a liability: a rare outcome sequence has small population, not zero, exactly as the rule requires.

Nothing in Sections 4–7 uses [T]. It is consumed exactly once, in Corollary 2, to pass from populations to observed frequencies; the value of the exponent is fixed by the ontology before any principle about credence is invoked.

3 Structural Assumptions

Three structural assumptions are used in the derivation. They are collected here, in one place, so that each result below can name exactly what it consumes. (Phase additivity is Lemma 1, not an assumption.)

Assumption 2 ([AS] Aggregate structure)**.** Class summaries are physical, and satisfy four clauses. *(Capacity)* The summary of a class is storable in a record of bounded capacity: its dimension is fixed, independent of the size of the class. *(Additivity)* The summary is an additive set function on sub-collections of a class: $\mathrm{summary}(E_1 \cup E_2) = \mathrm{summary}(E_1) + \mathrm{summary}(E_2)$ for disjoint E_1, E_2—one consistent bookkeeping across all coarse-grainings, whether or not each sub-collection is separately realizable as the class of a record. *(Marginality)* For independent subsystems, the joint summary is determined by the marginal summaries. *(Regularity)* The summary depends continuously on the phases it aggregates.

The clauses are named, rather than left implicit inside a lemma, because Lemma 3 uses each of them. Their motivation is supplied by the ontology: capacity is Postulate 2's lossiness (an unbounded summary, such as the full phase multiset, is not a record); additivity is the consistency of parent-level with class-level bookkeeping; marginality says that where no correlation exists, none may be consumed. Capacity, additivity, and marginality are thus physical clauses; regularity is the smoothing idealization, realized approximately by every finite model (Remark 5).

Assumption 3 ([LC] Linear composition, with attainability)**.** Preparation devices compose linearly on arrows (the coherent class summaries derived in Lemma 3 and named in Definition 4): a device carries the arrow vector $(\alpha_1, \ldots, \alpha_k)$, $k \geq 2$, to $U(\alpha_1, \ldots, \alpha_k)$ with U complex-linear and invertible.

Nontriviality (the empirical clause). Nontrivial devices exist: the admissible family includes maps that are not generalized permutations with unimodular entries (superposition-forming devices, e.g. beamsplitters of continuously variable angle). Physically, this clause is the assumption that interference is possible: a generalized permutation only relabels and re-phases outcome classes, leaving every population untouched, whereas a nontrivial device combines arrows from several classes into one, which is the operation that permits cancellation. A world admitting only such permutations exhibits no interference. The assumptions thus posit that interference *occurs*; what the theorems derive is its quantitative law. This clause is the paper's light postulate: a measured regularity taken as input, in the manner of a principle theory.

(At) Attainability (idealization clause, with invertibility above). The admissible family is rich: composing devices of continuously variable angle, the physically attained arrow magnitudes of single classes fill out an interval; magnitudes are *jointly* attainable on independent pairs of subsystems; and the attained arrow vectors are dense in direction. Exact density is an idealization that no single finite ontology satisfies—finitely many subsets yield finitely many attainable values—so (At) is to be read at $O(1/N)$ granularity, and the theorems hold correspondingly in robust form (Remark 5). (At) is the richness premise of the two theorems: Theorem 1 requires attained magnitudes to fill an interval jointly, so that the functional equation has a rich domain; Theorem 2 requires attained vectors dense in direction, so that a device conserving counts on all attainable preparations is, by continuity, a norm isometry.

Ontological reading. On Postulate 1, a device's action is a table of stage-relative phase increments and class recombinations. [LC] is then the statement that these increments are uniform at record level—they depend on the classes a configuration belongs to, never on configurations within a class. This is the device-level form of record-level indifference (Lemma 2), and it is what makes the device a matrix. [LC] remains an assumption; what the reading supplies is its meaning inside the ontology, not its derivation.

No norm-preservation is assumed; that devices preserve a norm, and which one, is derived (Theorem 2 and Remark 6).

Assumption 4 ([PR] Existence and universality of the population rule)**.** A population rule exists and is universal: a single strictly monotone increasing function f, common to all systems, preparations, and composites, relates populations to arrows via $n_i/N = f(|\alpha_i|)$ on the attainable magnitudes, where the arrows carry the parent-relative scale of Definition 4. That some lawful relation exists is the assumption that outcome statistics are reproducible and phase-dependent, both empirical facts. Two clauses carry weight in the proofs. *Universality* is not bookkeeping: it is what licenses applying the same f to subsystems and to their composites in Theorem 1, and applying it to every preparation a device can act on in Theorem 2. *Strict monotonicity* carries physical content as well as regularity: it reflects that $|\alpha_i|$ is population discounted by phase disagreement—more aligned phases, larger population—and it is the condition under which the multiplicative functional equation of Theorem 1 has only pure-power solutions [2].

3.1 Where the quantum structure enters

[LC] is where the linear structure of quantum mechanics enters this paper. Complex-valued class summaries composing under linear, invertible maps, with genuine superposition-forming elements—that is most of the kinematic skeleton of finite-dimensional quantum theory, and this paper assumes it rather than deriving it. What the theorems then add is deliberately narrow: *given* that structure, the exponent relating populations to arrows is forced to be 2 by counting alone, and the premise that forces it—conservation of a cardinality—is available only where classes have canonical measure (Proposition 1). The reader who wants linearity itself derived from deeper principles is directed to the reconstruction program [17, 8, 21]; the reader who wants to know what finiteness provides, over and above that program, is directed to Remark 7.

The net assumption count against the textbook axiom set, in one place: textbook quantum mechanics postulates linear composition, the Born rule, state normalization, and unitarity. This paper postulates linear composition ([LC], with its empirical core) and finiteness, and derives the Born exponent (Theorems 1–2), the normalization (Remark 6), and the norm-preservation of devices (Corollary 1) as counting identities, with the rule's domain fixed by stability (Theorem 4). Masanes, Galley and Müller [22] show the measurement postulates operationally redundant given the rest of quantum theory; the counting derivation locates *why*: they are consequences of conserved cardinality.

4 The Form of the Population Rule

Lemma 2 (Record-level indifference)**.** *Any observer-relative rule relating outcome classes can depend on a class only through record-level aggregates of the class.*

Proof. By Postulate 2, the members of a compatibility class are indistinguishable with respect to the record defining it. A rule whose inputs are record-level correlations has no variable with which to distinguish class members. □

Lemma 3 (The admissible aggregate is the coherent phase sum, through its magnitude)**.** *Under Postulates 1–2 and* [AS]*, the admissible class aggregate at a stage r is, up to a scale, the coherent sum $\sum_{\omega\in\Omega_{r_i}} e^{iS_r(\omega)}$, and any population rule depends on it only through its modulus.*

Proof. Four steps, each excluding a family of candidates. Throughout, phases are the stage-relative $S_r(\omega)$ of the stage whose partition is being described; the stage is fixed for the duration of the argument.

(i) Record level only. By Lemma 2, the rule cannot distinguish configurations within a class; its input can only be a symmetric summary of the class's phase multiset.

(ii) The summary is a per-configuration sum. By [AS]-additivity, the summary is an additive set function on sub-collections of the class; on a finite underlying set, an additive set function is a per-element sum, summary$(E) = \sum_{\omega\in E} g(\omega)$, and by step (i), g depends on ω only through its phase: summary$(E) = \sum_{\omega\in E} g(S_r(\omega))$. The rival summaries fall to the named clauses: the mean phase is not additive (the mean of a union is not the sum of the means); the full multiset *is* additive in the multiset sense but violates [AS]-capacity—its size grows with the class, and Postulate 2's lossiness excludes any summary that cannot live in a bounded record.

(iii) The summand is the unimodular exponential. It remains to fix g. By [AS]-marginality, for independent subsystems the joint summary is determined by the marginal summaries; by Lemma 1 the joint phases are the sums S_A+S_B. So the requirement reads: $\sum_{\omega_A}\sum_{\omega_B} g(S_A+S_B)$ is a function F of $\sum_{\omega_A} g(S_A)$ and $\sum_{\omega_B} g(S_B)$, for every pair of phase multisets. Singleton sub-collections give $g(x+y) = F(g(x), g(y))$; two-element sub-collections make F additive in each argument on attained values, and [AS]-regularity extends this to real-biadditivity, whose continuous solutions are $F(u,v) = a\,uv + b\,u\bar{v} + c\,\bar{u}v + d\,\bar{u}\bar{v}$. Symmetry of $x+y$ forces $b = c$, and requiring F to compose associatively across three independent subsystems—apply F in the two bracketings to singleton triples with generic phases—eliminates every conjugate-bearing term, leaving $F(u,v) = a\,uv$ with a absorbable into the scale of g: hence $g(x+y) = g(x)\,g(y)$. The continuous solutions on angles (phases are defined modulo 2π) are the characters of the circle: $g(S) = e^{imS}$ with $m \in \mathbb{Z}$. The case $m = 0$ degenerates the summary to the bare cardinality, making every statistic phase-independent, contradicting the empirical anchor of [PR] (statistics are phase-dependent; interference occurs). The sign of m is a convention (complex conjugation of every summary), and $|m| > 1$ merely relabels the phase variable: only mS ever enters a record-level quantity, so mS *is* the physical phase, and we name it S. The canonical summand is therefore $e^{iS_r(\omega)}$, and the admissible aggregate is the coherent sum $\sum_{\omega\in\Omega_{r_i}} e^{iS_r(\omega)}$, up to a scale. Note that the additivity of step (ii) holds for the complex sums themselves, not for their moduli: $|\alpha_1+\alpha_2| \le |\alpha_1| + |\alpha_2|$ in general, and this gap is precisely interference (Corollary 3).

(iv) Only its modulus. Within a stage, only phase differences are physical: the zero point of S_r is a convention of description, not a fact. Consider one class described twice: once with phases $(90°, 60°, 30°)$, and once with every phase at the stage shifted by $30°$, giving $(120°, 90°, 60°)$. Every pairwise phase difference is identical; every correlation, record, and class is identical; these are one stage under two conventions. Yet the coherent sum's direction differs between the descriptions ($60°$ versus $90°$) while its modulus does not. A population is a fact about the world—the cardinality of one fixed set—and cannot differ between conventional redescriptions. Dependence on the argument of the sum is therefore excluded, and the rule may consume only the modulus. (Direction is not thereby idle: relative angles between the sums of distinct classes are convention-independent and act upstream, where devices combine them; the population rule reads the moduli that result.) □

Definition 4 (Arrow, with parent-relative scale)**.** For an outcome class r_i of a partition of the parent class Ω_r at a stage, write $c_i = \sum_{\omega\in\Omega_{r_i}} e^{iS_r(\omega)}$ for the raw coherent sum. The *arrow* of the class is

$$\alpha_i = \frac{c_i}{A_r}, \tag{3}$$

where $A_r > 0$ is a scale associated with the parent class at that stage—one scale per parent, common to all classes of all partitions of that parent. The scale is not a free constant: its value is determined by the theory in Remark 6, and its parent-relativity is not optional—a single scale serving two nesting levels at once is inconsistent (apply [PR] with one fixed scale both to a partition of Ω_r and to the sub-partition of one of its cells, and the two applications force the cell to exhaust its parent). For independent subsystems the scales compose multiplicatively, $A_{AB} = A_A A_B$, consistently with Lemma 4. This definition introduces *notation only*: the form of the aggregate was derived above.

By Lemmas 2–3 and [PR], the population rule is of the form

$$\frac{n_i}{N} = f(|\alpha_i|), \tag{4}$$

with f a universal strictly monotone function on the attainable magnitudes, otherwise undetermined. The remainder of the derivation determines f.

Lemma 4 (Factorization)**.** *For independent subsystems: $c_{(i,j)} = c_i\, c_j$ and hence $\alpha_{(i,j)} = \alpha_i \alpha_j$ under the composed scale $A_{AB} = A_A A_B$; $|\Omega_{(i,j)}| = |\Omega_{r_i}| \cdot |\Omega_{r_j}|$; and the joint parent class population is the product of the marginal parent populations, $N_{AB} = N_A N_B$.*

Before the proof, the terms. A *subsystem* is a subset of the degrees of freedom—part of what every configuration specifies, not a subset of the configurations. Two subsystems A, B are *independent*, relative to a class, when the class exhibits full crossing of their values: every occurring arrangement of A's degrees of freedom co-occurs with every occurring arrangement of B's, in product proportion, with each configuration's phase splitting additively across the two (Lemma 1). Independence is thus a pattern in the census of a class, not a relation between individual configurations. Two distinct composition operations now coexist and must not be conflated: within one partition, disjoint classes' populations *add* (the arithmetic of Section 5); across independent subsystems, populations *multiply* (the arithmetic of this section). Correlated subsystems—entangled classes, where combinations are missing—violate the crossing premise, and the lemma correctly does not apply to them. One existence premise is therefore carried explicitly into Theorem 1: *some independent pair of subsystems exists, and by (At) the required magnitudes are jointly attainable on it.* A product-structured ontology supplies the pair in any two uncorrelated systems, and the explicit model of Section 8 realizes it by construction.

Proof. The joint class is $\Omega_{r_i} \times \Omega_{r_j}$: by full crossing, its members are exactly the pairs, so its cardinality is the product of the cardinalities; the same applies to the joint parent class, giving $N_{AB} = N_A N_B$. For the raw sums: by Lemma 1, $\sum_{\omega_A} \sum_{\omega_B} e^{i(S(\omega_A)+S(\omega_B))} = \big(\sum_{\omega_A} e^{iS(\omega_A)}\big)\big(\sum_{\omega_B} e^{iS(\omega_B)}\big)$; dividing by $A_{AB} = A_A A_B$ gives $\alpha_{(i,j)} = \alpha_i \alpha_j$. □

Theorem 1 (Power-law form)**.** *Under Lemma 1,* [AS], [PR]*, and the attainability clause* (At) *of* [LC]*, together with the existence of an independent pair of subsystems: $f(x) = x^p$ for some $p > 0$, on the attainable magnitudes.*

Proof. Suppose, for contradiction, that f is not a pure power on the attainable set M of arrow magnitudes. By (At), M contains an interval with joint attainability on an independent pair, and by Lemma 4 it is closed under products of attainable pairs. A strictly monotone function satisfying $f(xy) = f(x)f(y)$ on a set containing an interval is necessarily a pure power: setting $g(u) = \log f(e^u)$ converts the equation to Cauchy's additive equation, whose monotone solutions are linear [2], giving $f(x) = x^p$; strict monotonicity forces $p > 0$, and $f(1) = 1$ follows. Hence if f is not a pure power, there exist *jointly attainable* x_0, y_0 with $f(x_0 y_0) \neq f(x_0)f(y_0)$. Realize x_0

and y_0 as arrow magnitudes of outcome classes of two independent subsystems—(At) provides exactly this. Now compute the population fraction of the joint class twice. Directly: the joint class has arrow magnitude $x_0 y_0$ (Lemma 4, arrows multiply under the composed scale) and parent population $N_A N_B$ (Lemma 4, parents multiply), so by [PR], applied to the composite as universality licenses,

$$\frac{n_{(i,j)}}{N_A N_B} = f(x_0 y_0).$$

As a product of marginals: the joint class population is the product of the marginal class populations (Lemma 4, classes multiply), so

$$\frac{n_{(i,j)}}{N_A N_B} = \frac{n_i}{N_A} \cdot \frac{n_j}{N_B} = f(x_0)\, f(y_0).$$

The joint class is one fixed set with one cardinality; the two computations must agree; they do not. Contradiction. Therefore $f(x) = x^p$ with $p > 0$. $\square$

5 The Exponent

First, the ontological reading of a device, on which the conservation premise rests. A device stands between two stages of an observer chain: the upstream stage bears the preparation record; the downstream stage bears the same record together with passage correlations; and *neither bears outcome information.* The two stages' records therefore constrain configurations identically—no configuration compatible with the preparation is excluded by unread passage—so the two stages share one compatibility class: the same fixed set of N configurations, partitioned upstream one way and downstream another. What differs between the stages is not the set but the phase assignment: $S_{r'}(\omega) = S_r(\omega) + \Delta S(\omega)$, with the increments record-level uniform by [LC]'s ontological reading. On a context-free picture this same-set reading is inconsistent (Remark 4); on the relational picture it is exact, and it makes conservation automatic:

$$\sum_i n_i = N \quad \text{for every preparation, upstream and downstream of every device,} \tag{5}$$

because re-partitioning a fixed finite set cannot change its cardinality. Since upstream and downstream partitions share the parent, they share the parent-relative scale A_r; substituting the rule $n_i = N f(|\alpha_i|) = N|\alpha_i|^p$ from Theorem 1 into (5) on both sides gives, for every preparation,

$$\sum_i |\alpha_i^{\text{up}}|^p \;=\; 1 \;=\; \sum_j |\alpha_j^{\text{down}}|^p. \tag{6}$$

Two readings of (6). First, the constraint is state-independent because the rule references only the class and its parent—never the global state—as additivity across independent subsystems requires. Second, since $\alpha^{\text{down}} = U\alpha^{\text{up}}$ by [LC], every admissible device carries arrow vectors of unit ℓ^p norm to arrow vectors of unit ℓ^p norm, where $\|\alpha\|_p = (\sum_i |\alpha_i|^p)^{1/p}$, of which the $p = 2$ case is ordinary Pythagorean length. By (At) the attainable vectors are dense in direction, so by continuity and homogeneity every admissible device preserves the ℓ^p norm of every vector: it is a linear ℓ^p isometry. The physical requirement—counting conservation—has become a classified mathematical condition.

Theorem 2 (Exponent)**.** *Under* (5)*,* [PR]*, and* [LC]*, the exponent of Theorem 1 is* $p = 2$*.*

Proof. Suppose, for contradiction, that some $p \neq 2$ is admissible. By the argument above, every admissible device is a complex-linear isometry of the ℓ^p norm on $\mathbb{C}^k$, $k \geq 2$. For $p \neq 2$, the only such isometries are generalized permutations—permutation matrices composed with diagonal unimodular factors [20]. But [LC] asserts the existence of devices that are not generalized permutations. Contradiction. Therefore $p = 2$. Sufficiency: for $p = 2$, conservation holds for every U preserving $\sum_i |\alpha_i|^2$, a family containing continuous rotations of all angles; for illustration, the device carrying $(1, 0)$ to $(\cos\theta, \sin\theta)$ conserves $\cos^2\theta + \sin^2\theta = 1$ identically in θ, whereas for $p \neq 2$ even this single family fails ($2^{1-p/2} \neq 1$ at $\theta = \pi/4$). □

Remark 5 (Robust form in a literally finite ontology)*.* In any single finite ontology the attainable magnitudes and directions are finite sets, and populations are integers, so (4) and (6) hold to $O(1/N)$ rather than exactly. The theorems survive in robust form: the multiplicative Cauchy equation is stable—functions satisfying it approximately are approximated by exact pure powers [19]—and approximate ℓ^p isometries on ϵ-dense sets are near exact ones, so a rule within $O(1/N)$ of the constraints is within $O(1/N)$ of x^2. The idealized statements are the $N \to \infty$ limits; the enumeration of Section 8 exhibits the $O(1/N)$ residues directly, with the measured $1/N$ scaling as the empirical face of the robustness.

Corollary 1 (Devices are 2-norm isometries)**.** *Under* (5)*, every admissible device preserves* $\sum_i |\alpha_i|^2$*. Norm-preservation is thus a consequence of counting conservation, not an assumption; the theorem's content may be stated: a finite counting ontology with linear devices either admits no superposition-forming devices at all, or obeys the Born exponent.*

Remark 6 (Normalization is forced, not chosen)*.* With $p = 2$, equation (6) reads $\sum_i |\alpha_i|^2 = 1$ for every partition of every parent. This is not a convention adopted for convenience: it is the counting identity $\sum_i n_i = N$ expressed through the derived rule, and it fixes the parent-relative scale, $A_r^2 = \sum_i |c_i|^2$—a value the theory itself then proves is *partition-invariant* within each parent, since every stable partition of the parent must yield the same N. The familiar unit-norm condition on quantum states appears here as a theorem about cardinality—the statement that the configurations of a class are all still there after any re-partitioning—rather than as a postulate about state vectors.

Remark 7 (On the classical core of the argument)*.* The mathematics of Theorem 2 is elementary and classical: rotations preserve sums of squares and no other power sums, and conservation under linear evolution is known to select the 2-norm [1, 20]. Those arguments are temporal: a norm is preserved under evolution in time. Here nothing evolves—the two stages a device connects are positions in a refinement ordering (Remark 3), and the derivation is the timeless form of the same selection, with "before and after the device" replaced by "coarser and finer records of one fixed set." No mathematical novelty is claimed, here or anywhere in this paper. What the present framework contributes is the premise. In an ontology without canonically measurable classes, "the total amount of configuration space" is defined only relative to a chosen measure, and conservation of a convention forces nothing; the arguments cited above accordingly conserve *probability*, a normalization convention about descriptions. Here the conserved quantity is a cardinality—a number of things—whose conservation is a fact of the ontology (Postulate 1 and the same-set reading of devices), prior to any probabilistic notion. Proposition 1 states the point in general form.

Proposition 1 (Necessity: canonically measurable classes)**.** *The Born exponent can be derived, rather than decreed, only in an ontology whose outcome classes carry a canonical measure:*

one available without further choice. Finite cardinality is the unique assumption-free case; and since the derivation consumes only $N = |\Omega_r|$—the total $|\Omega|$ never enters—what an ontic Born exponent dictates is the hereditary finiteness of compatibility classes. Global finiteness of Ω is the parsimonious way to secure it, and is independently motivated, but is not forced by this argument alone.

Argument. "The fraction of configurations in an outcome class" requires a measure. For infinite classes, cardinality fails—countably infinite sets admit bijections with proper subsets, so ratios of infinite counts have no meaning—and continua quantify size only relative to a measure $d\mu$, whose choice is mathematically equivalent to choosing the probability rule. An infinite ontology with sufficient symmetry does carry an invariant measure (normalized Haar measure on a compact homogeneous space; the Fubini–Study measure on projective Hilbert space), but the symmetry structure is itself an ontological posit, and choosing it is choosing the measure: the conventionality is relocated, not removed. Positing finiteness is also a posit; the asymmetry claimed is comparative, not absolute—cardinality is the only measure that comes with *no further* structure beyond the set itself, whereas every infinite alternative requires selecting structure whose selection is measure-equivalent. A continuum can host the quadratic measure as a postulate—Bohmian quantum equilibrium does so [11]—but cannot force it. □

6 The Born Theorem

Theorem 3 (Born by counting). *Under Postulates 1–3 and* [AS], [LC], [PR], *with an independent pair of subsystems attainable:* $n_i = N|\alpha_i|^2$, *and*

$$P(r_i \mid r) = \frac{|\Omega_{r_i}|}{|\Omega_r|} = |\alpha_i|^2$$

to $O(1/N)$ integer rounding (Remark 5). Assumption [T] *is not consumed.*

Corollary 2 (Frequencies; consumes [T]). *Consider k repeated runs of a preparation along one observer chain, under two conditions carried by [PR]'s universality: each run is reproducible (the same preparation, hence the same arrows, at each stage), and successive runs are independent subsystems of the record-bearing composite, so that Lemma 4 iterates across runs. Then the sequence classes carry product populations, and counting the partition by stored k-run outcome frequency is a binomial computation: the population concentrates on stored counts whose frequencies lie near $|\alpha_i|^2$, with the usual $O(1/\sqrt{k})$ width. By* [T], *the typical records are the ones to reason from; hence the outcome frequencies observers record track $|\alpha_i|^2$. This is the Born rule as an empirical law, and it is exactly one assumption away from the population theorem. One finite-ontology caveat should be noted: sequence-class populations scale as $N\prod_i p_i$, so frequency reasoning by literal counting is supported for run lengths $k \lesssim \log N$—a finite world underwrites frequency talk only over runs shorter than its logarithm, which is itself a prediction of the framework.*

Corollary 3 (Interference and decoherence as arithmetic). *The coherent sums are additive over disjoint unions of classes—$c_{1\cup 2} = c_1 + c_2$, since every phase is counted once either way—but their moduli are not:*

$$|\alpha_1 + \alpha_2| \;\leq\; |\alpha_1| + |\alpha_2|,$$

with equality only when the arrows are aligned. Populations read squared moduli, so a parent class counted whole and the same class counted as two children generally disagree:

$$|\alpha_1 + \alpha_2|^2 = |\alpha_1|^2 + |\alpha_2|^2 + 2\,\mathrm{Re}(\alpha_1^* \alpha_2),$$

and the cross term—the gap between the modulus of the sum and the sum of the moduli—is interference. Which counting is physical is fixed by the record structure: arrows sum only within a compatibility class, so if no record distinguishes the children they form one class and the parent-level count (with cross term) governs, whereas a which-path record makes the children distinct classes, forces separate summation, and removes the cross term. That removal is decoherence. Conditionalization is counting over surviving configurations; no collapse postulate is required. Total probability is unity by finite additivity of cardinality.

Example (The two-slit experiment as counting). Corollary 3 contains the two-slit experiment. With no which-path apparatus, no degree of freedom in any configuration is correlated with slit-membership: slit-sector is sub-record structure, invisible to every partition an observer can hold, so the configurations compatible with the observer's records form one class spanning both sectors. The screen-position refinement sums over that whole class; the cross term is present; the populations across screen positions carry fringes. With a which-path apparatus present, every configuration contains a register degree of freedom correlated with slit-sector: the correlation exists in the configurations themselves, so slit-membership is record-level structure, and the class is—at the level any record can access—already two disjoint classes. Every downstream sum proceeds per class; the cross term never forms; the populations are fringeless. Nothing destroys the interference: the cross term is simply unlicensed, because arrows sum only within compatibility classes and the correlation structure of the configurations fixes what the compatibility classes are. Counting is therefore not invariant under the insertion of intermediate record-level refinements—where refinement occurs in the ordering determines which sums are taken—and this placement-dependence is the counting form of which-path complementarity. Which partitions exist at record level is exactly what stability formalizes (Section 7).

7 Domain: Stability

Definition 5 (Recordability, redundancy, stability)**.** A partition is recordable if correlating a fresh register with it leaves all downstream observer statistics unchanged. Its redundancy capacity is the number of independent copies that can be so correlated before downstream statistics degrade. A record is stable if recordable with redundancy up to the ontology's capacity. All three notions compare counts across positions in the refinement ordering; no dynamics is invoked.

Theorem 4 (Domain of the rule)**.** *Stable partitions are exactly those whose classes carry no downstream cross-interference. On stable partitions the record-level description closes and Theorem 3 applies. Unstable partitions support no persistent records and hence no observer-relative outcome facts.*

Proof sketch. If a partition's classes carry no cross terms, arrows recombine additively across it, so a correlated register changes no downstream count, and the argument iterates for as many registers as the ontology's capacity supports. Conversely, cross terms present: a correlated register renders the classes disjoint and removes the cross terms, altering downstream counts

(Corollary 3); each further copy compounds the alteration. Quantitatively: k partial records at coupling strength χ leave fringe visibility

$$V = (\cos\chi)^k,$$

so any nonzero coupling degrades an unstable partition's statistics geometrically in the number of copies, while a stable partition (χ effectively zero in the aligned basis) sustains redundancy to capacity; and visibility and distinguishability obtained from integer counts obey

$$V^2 + D^2 = 1.$$

Both relations are verified by literal enumeration in Section 8. □

Three consequences follow and are verified below: the zero-disturbance recording basis is selected by the criterion [30]; the geometric redundancy decay $(\cos\chi)^k$ recovers the redundancy structure of quantum Darwinism [32]; and the duality relation $V^2+D^2 = 1$ recovers the visibility–distinguishability tradeoff [29, 13] from integer counts.

8 Computational Verification

An explicit finite model verifies every theorem by literal enumeration: configurations are tuples (sector, micro-id, phase-index, record) with phases on a lattice of 3,600 values and class populations N up to 4.2×10^6; probabilities are computed by materializing configuration lists and counting elements; the counting rule was fixed before any case was computed; preparation parameters enter only through correlation structure. The model's phase bookkeeping is stage-relative—class phases are re-derived at each stage of the chain from the device's increments, exactly as Postulate 1 prescribes, and Remark 4 is the reason it must. The model serves as an existence proof for the postulates and, at $O(1/N)$ granularity, for the assumptions: an explicit structure satisfying Postulates 1–2 exactly and realizing [AS], [LC], [PR] to the tolerances of Remark 5, establishing consistency and non-vacuity of the axiom set in the robust sense.

- Unequal splits: counted P matches $\cos^2\theta$ exactly for dyadic values and to $\sim 3 \times 10^{-7}$ otherwise ($\theta = 15°$: 933,013 of 10^6 configurations against $\cos^2 15° = 0.9330127$); three-outcome and biased three-outcome splitters, errors $\leq 3 \times 10^{-7}$.
- Interference and decoherence: Mach–Zehnder fringe visibility 1.0000 by counting, tracking $\cos^2(\delta/2)$; a which-path record yields visibility 0.0000.
- Conditionalization: a refinement eliminates 250,000 of 10^6 configurations; a second device counted over the survivors reproduces the conditional Born value to 5×10^{-7}.
- Exponent discrimination: with the scale calibrated once and held fixed—the model's implementation of the state-independence that Section 5 requires—count conservation is violated by 15–41% for $p \in \{1, 1.5, 2.5, 3\}$ and satisfied exactly at $p = 2$; product consistency is violated at the 10^{-3}–10^{-2} level by monotone non-power rules and satisfied at rounding level ($\lesssim 10^{-6}$) by every pure power. The two constraints separate exactly as the two theorems separate: products eliminate non-powers (Theorem 1), conservation eliminates $p \neq 2$ (Theorem 2).

- Entanglement: multiplicativity fails for a Bell-type state, as required (Lemma 4 presupposes independence); the entangled partner acts as a which-path record (unconditional visibility 0.0000), and conjugate-basis measurement restores conditional fringes at visibility 1.0000 with opposite phases. The full CHSH treatment is the second companion paper [25].

- Stability: zero-disturbance recording at the bundle-aligned basis; redundancy decay matching $(\cos\chi)^k$ at every grid point; $V^2 + D^2 = 1$ with maximum deviation 3×10^{-4}.

- Rational residue: $|P - \cos^2\theta|$ scales as $O(1/N)$; measured log–log slope -1.06—the empirical face of Remark 5.

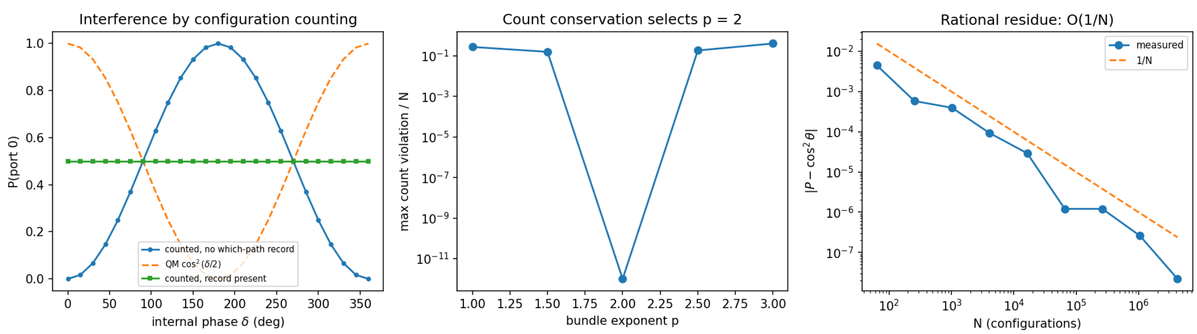

Figure 1: Interference by counting, with and without a which-path record (left); count conservation selecting $p = 2$ (center); $O(1/N)$ residue (right).

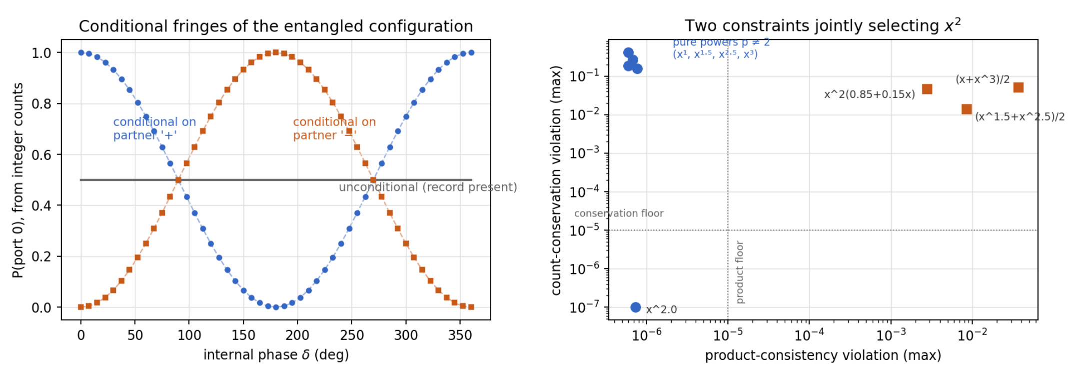

Figure 2: Conditional fringes of the entangled configuration (left); product consistency and conservation jointly selecting x^2 (right).

9 Status of Every Claim

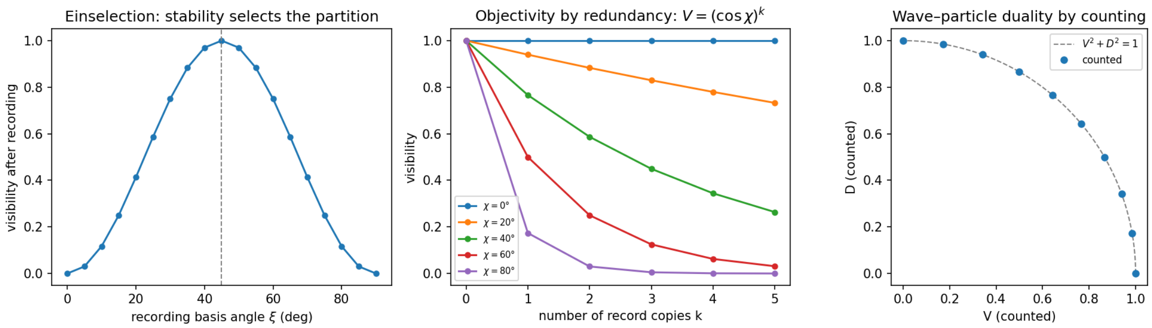

Figure 3: Stability selects the recordable partition (left); redundancy decay $V = (\cos\chi)^k$ (center); $V^2 + D^2 = 1$ from integer counts (right).

Item	Status	Where
Finite timeless Ω; two-level relational phase	Postulate	Post. 1
No context-free phase assignment possible	Proved	Rem. 4
Records and compatibility classes	Postulate	Post. 2
Observers structural; observer chains	Definition (vocabulary only)	Def. 1
Existence of observer chains	Postulate	Post. 3
Sequential language = refinement order	Convention, anchored once	Rem. 3
Probability as counting	Definition	Def. 2
Probability = population (functional role)	Identification	§2.2
[T] typicality	Assumption (frequency reading only)	Ass. 1
Phase additivity	Proved (from Post. 1)	Lem. 1
[AS] aggregate structure	Assumption (four named clauses)	Ass. 2
[LC] linear composition + (At)	Assumption (ontological reading given)	Ass. 3
[PR] existence/universality of the rule	Assumption (empirically anchored)	Ass. 4
Indifference; coherent sum; magnitude-only	Proved under [AS]	Lems. 2–3
Parent-relative scale (global scale inconsistent)	Proved	Def. 4
Factorization (independent case)	Proved	Lem. 4
Power-law form	Proved by contradiction	Thm. 1
Same-set conservation across devices	Proved (relational phase)	§5
Exponent $p = 2$	Proved by contradiction	Thm. 2
Robust form at finite N	Argued (stability of the equations)	Rem. 5
Normalization $\sum \lvert\alpha_i\rvert^2 = 1$ per parent	Derived (forced, not conventional)	Rem. 6
Classical core of Thm. 2	Known; cited	Rem. 7
Necessity: finite classes	Argued (Haar caveat conceded)	Prop. 1
Born by counting	Proved	Thm. 3
Frequencies track $\lvert\alpha_i\rvert^2$	Argued + [T] ($k \lesssim \log N$)	Cor. 2
Interference/decoherence as arithmetic	Proved	Cor. 3
Domain via stability	Proved (sketch) + verified	Thm. 4
Postulates exactly; assumptions to $O(1/N)$	Verified by enumeration	§8
Base-level phase texture	Open problem	§11
Origin of [LC], [AS]	Open problems	§3.1, §11
$O(1/N)$ deviation from exact Born	Prediction	below

Predictions. With finite N, every probability is m/N; Born statistics hold to accuracy $O(1/N)$—empirically invisible for large N, falsifiable in principle, and a definite point of difference from standard quantum mechanics. Finiteness likewise implies a maximum entropy $\log N$ per class, and Corollary 2 adds a third: literal-counting frequency reasoning is supported only over run lengths $k \lesssim \log N$. One refinement of the first prediction: the deviation scale for a given experiment is set not by any global count but by the unresolved multiplicity of the measured sector itself—under independence, environment multiplicity multiplies every outcome class and the parent alike and cancels in the ratios—so the deviation grows as refinement penetrates the measured sector, a graded form of these predictions developed, with its falsifiability accounting, in the framework paper [26].

10 Relation to Prior Work

Branch counting fails in Everettian quantum mechanics [15] because branch number is coarse-graining-relative and undefined [28]; configurations, by contrast, are fundamental entities with a canonical cardinality. Decision theory [9, 27], envariance [31], and self-locating uncertainty [24] obtain Born weights from principles about credences; the present derivation obtains the populations from ontology, with the compatibility class serving as an ontic count rather than a self-location credence set, and confines the credence-facing residue to the single assumption [T]. It is, to our knowledge, the only *mechanical* account of quantum probability on offer: the alternatives either derive credences or postulate the measure they host, whereas here probability is frequency by literal cardinality over equally real entities. The finite-dimensionality this requires has independent advocates [7, 5, 3]. Hardy's excess-baggage theorem [18] shows that ontological models reproducing even qubit statistics require an infinite ontic state space; it does not apply here, because its framework assigns each ontic state response probabilities for every measurement—precisely the counterfactual values this ontology denies: a configuration bears outcome records only for the devices its correlation structure contains, the same structural fact that carries the Bell analysis of the second companion paper [25]. Gleason-type theorems [14, 6] force the quadratic measure given Hilbert-space structure; Masanes, Galley and Müller [22] show the measurement postulates operationally redundant given the remainder of quantum theory; the relation of Theorem 2 to the known 2-norm selection results is stated in Remark 7. Hall, Deckert and Wiseman [16] reproduce quantum phenomena with a finite ensemble of classical worlds whose density is driven to $|\psi|^2$ dynamically; the stability criterion of Section 7 plays the corresponding selecting role statically, while the origin of the linear structure remains open in both. Bohmian mechanics [11] postulates the $|\psi|^2$ measure over a continuum, consistent with Proposition 1. The background ontology belongs to the timeless tradition of Page and Wootters [23], Barbour [4], and canonical formulations [10]. The stability results express einselection [30], redundancy [32], and the duality relation [29, 13] in counting form; the phenomena are standard, and the reformulation and the domain theorem are the contribution claimed.

11 Limitations and Conclusion

The conditions stated in Section 1 govern all claims, and four further limits are stated here explicitly.

Phase texture. The theorems constrain the joint assignment of populations and stage-relative phases: not every mathematically describable assignment satisfies the derived relations, and the theorems entail that such assignments are not realized. The relational phase structure of Postulate 1 removes the overdetermination that a context-free assignment would suffer (Remark 4); what remains open is the base-level mechanism by which admissible assignments are populated—in particular, that generic micro-textures satisfy the relations typically, and what determines the increments ΔS that [LC] organizes.

Mathematical novelty. None is claimed. The core of Theorem 2 is the invariance of the 2-norm under rotations and the absence of nontrivial ℓ^p isometries for $p \neq 2$ [20]; the observation that conservation plus linearity selects the 2-norm has been made before [1]. The contribution is that the conservation premise is here an ontological fact—a cardinality of one fixed set shared by the stages a device connects—rather than a normalization convention, together with the necessity argument (Proposition 1, Remark 7).

The seat of linearity. As Section 3.1 states, [LC] imports the linear structure of quantum mechanics; this paper derives the exponent given that structure, not the structure itself. [AS] likewise names, rather than derives, the aggregate structure the coherent sum requires.

Idealization. The exact theorems live at $N \to \infty$; every finite ontology realizes them to $O(1/N)$ (Remark 5).

Within these limits, the result is: in a finite, timeless configuration ontology with relational phase, the population of every stable outcome class is $N|\alpha|^2$—the power-law form forced by the multiplicativity of counting, the exponent forced by conservation of the number of configurations, the normalization forced rather than chosen, and the domain fixed by the stability of records. Probability, defined as counting, thereby obeys the Born rule; observed frequencies track it given one stated typicality assumption, on which the exponent does not depend.

The argument may also be run in reverse. If the Born exponent is a fact about the world rather than a principle about credences, then the world's compatibility classes admit only finitely many configurations: one may retain infinite classes or an ontic Born exponent, but not both, and a globally finite configuration space is the parsimonious ontology that secures the former. Read this way, the result joins a family of inferences in which a well-confirmed law, taken ontically, dictates a feature of fundamental ontology that the law's surface form conceals: relativistic simultaneity, read ontically, yields the block universe; the Wheeler–DeWitt constraint, read at face value, eliminates fundamental time [10]; fluctuation phenomena, read as facts about matter, yielded atoms. The present inference is of the same kind: the quadratic law of quantum probability, taken as a fact, yields a static ontology of countable classes—and the derivation above exhibits the mechanism, since it is precisely the conserved cardinality of such classes that forces the quadratic form.

Acknowledgments

The author used large language models (OpenAI ChatGPT; Anthropic Claude) to assist in the preparation of this work. The author takes full responsibility for all content.

References

[1] S. Aaronson, "Is Quantum Mechanics An Island In Theoryspace?" arXiv:quant-ph/0401062 (2004).

[2] J. Aczél, *Lectures on Functional Equations and Their Applications* (Academic Press, 1966).

[3] T. Banks, "Cosmological breaking of supersymmetry?" Int. J. Mod. Phys. A **16**, 910 (2001).

[4] J. Barbour, *The End of Time* (Oxford University Press, 1999).

[5] J. D. Bekenstein, “Universal upper bound on the entropy-to-energy ratio for bounded systems,” Phys. Rev. D **23**, 287 (1981).

[6] P. Busch, “Quantum states and generalized observables: a simple proof of Gleason’s theorem,” Phys. Rev. Lett. **91**, 120403 (2003).

[7] S. M. Carroll and A. Singh, “Mad-Dog Everettianism: Quantum Mechanics at Its Most Minimal,” arXiv:1801.08132 (2018).

[8] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Informational derivation of quantum theory,” Phys. Rev. A **84**, 012311 (2011).

[9] D. Deutsch, “Quantum theory of probability and decisions,” Proc. R. Soc. Lond. A **455**, 3129 (1999).

[10] B. S. DeWitt, “Quantum theory of gravity. I. The canonical theory,” Phys. Rev. **160**, 1113 (1967).

[11] D. Dürr, S. Goldstein, and N. Zanghì, “Quantum equilibrium and the origin of absolute uncertainty,” J. Stat. Phys. **67**, 843 (1992).

[12] A. Einstein, “What is the theory of relativity?,” *The Times* (London), 28 November 1919; reprinted in *Ideas and Opinions* (Crown, 1954).

[13] B.-G. Englert, “Fringe visibility and which-way information: an inequality,” Phys. Rev. Lett. **77**, 2154 (1996).

[14] A. M. Gleason, “Measures on the closed subspaces of a Hilbert space,” J. Math. Mech. **6**, 885 (1957).

[15] N. Graham, “The measurement of relative frequency,” in B. S. DeWitt and N. Graham (eds.), *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton University Press, 1973).

[16] M. J. W. Hall, D.-A. Deckert, and H. M. Wiseman, “Quantum phenomena modelled by interactions between many classical worlds,” Phys. Rev. X **4**, 041013 (2014).

[17] L. Hardy, “Quantum theory from five reasonable axioms,” arXiv:quant-ph/0101012 (2001).

[18] L. Hardy, “Quantum ontological excess baggage,” Stud. Hist. Phil. Mod. Phys. **35**, 267 (2004).

[19] D. H. Hyers, “On the stability of the linear functional equation,” Proc. Natl. Acad. Sci. USA **27**, 222 (1941).

[20] J. Lamperti, “On the isometries of certain function-spaces,” Pacific J. Math. **8**, 459 (1958).

[21] L. Masanes and M. P. Müller, “A derivation of quantum theory from physical requirements,” New J. Phys. **13**, 063001 (2011).

[22] L. Masanes, T. D. Galley, and M. P. Müller, “The measurement postulates of quantum mechanics are operationally redundant,” Nat. Commun. **10**, 1361 (2019).

[23] D. N. Page and W. K. Wootters, “Evolution without evolution: Dynamics described by stationary observables,” Phys. Rev. D **27**, 2885 (1983).

[24] C. T. Sebens and S. M. Carroll, “Self-locating uncertainty and the origin of probability in Everettian quantum mechanics,” Brit. J. Phil. Sci. **69**, 25 (2018).

[25] R. Tillitt, “CHSH by Counting: Bell Violation in a Finite Configuration Ontology,” companion paper, Zenodo (2026), `doi:10.5281/zenodo.21523901`.

[26] R. Tillitt, “Configuration Realism: Foundations,” companion paper, Zenodo (2026), `doi:10.5281/zenodo.21523984`.

[27] D. Wallace, “Quantum probability from subjective likelihood: improving on Deutsch’s proof of the probability rule,” Stud. Hist. Phil. Mod. Phys. **38**, 311 (2007).

[28] D. Wallace, *The Emergent Multiverse* (Oxford University Press, 2012).

[29] W. K. Wootters and W. H. Zurek, “Complementarity in the double-slit experiment,” Phys. Rev. D **19**, 473 (1979).

[30] W. H. Zurek, “Decoherence, einselection, and the quantum origins of the classical,” Rev. Mod. Phys. **75**, 715 (2003).

[31] W. H. Zurek, “Probabilities from entanglement, Born’s rule from envariance,” Phys. Rev. A **71**, 052105 (2005).

[32] W. H. Zurek, “Quantum Darwinism,” Nat. Phys. **5**, 181 (2009).

Code and data: `cr_toy_model.py`, `step1_multiplicativity.py`, `step4_stability.py`, `chsh_by_counting.py` (available with this paper).