

Observational Equivalence and Physical Distinguishability

Part 10 of *The Classification of Mathematical Representations of Physical Systems*

Matthew Long

The YonedaAI Collaboration, YonedaAI Research Collective

Chicago, Illinois, USA

m.long@yonedalabs.ai

<https://yonedalabs.ai>

October 2026

Abstract

The classification problem of mathematical representations of physical systems is a quotient: classify objects of a category of systems up to an equivalence relation. The quotient is unusable until the relation is chosen, and the choice is decisive. This paper develops the observational side of that choice. It separates three relations: mathematical equivalence, given by isomorphism of the component data of a representation; physical equivalence, an open choice constrained by what the identifying morphism preserves; and observational equivalence, induced by a functor from systems to observation records. Observation functors are defined, their induced relations are proved to be equivalence relations with classes given by record isomorphism classes, and their interaction with coarse-graining is proved monotone in one direction only. Distinguishability is treated through monotone metrics: the Fisher metric is the unique classical choice up to constant, while the quantum case admits a parametric family indexed by operator monotone functions. An explicit two-dimensional computation shows the resulting quadratic forms differ on non-commuting directions, so quantum metric distinguishability is not canonical. The plain tensor-category criterion for physical equivalence fails outright: equivalence of representation categories without the symmetric structure does not recover the group, and every proposed relation must name the layers of structure it preserves. A counterexample search records which candidate criteria fail and why, and the empirical layer of the three-layer architecture is instantiated as a functor and left open where the sources leave it open.

1 Introduction

The program of this series asks whether mathematical representations of physical systems admit a structural classification analogous to the classification of finite simple groups. The question is stated as

$$\text{Classify}(\text{Ob}(\text{Phys})/\simeq),$$

and each term of that expression requires work. The objects are tuples of state space, observable algebra, dynamics, symmetry data, and composition rules. The category they live in is still a schema: the preservation laws and composition rules for its morphisms are not fixed. The relation $\simeq$ is a choice, and the choice is essential. Two mathematical descriptions may be physically equivalent without being mathematically isomorphic; dual theories may describe the same observable physics

through very different mathematical structures.¹

This paper develops the empirical side of that quotient. Observational equivalence identifies two systems exactly when no admissible observation separates them. The formal object is the observation functor named by the Information Geometry Correspondence program of the YonedaAI research summary; that program names the object without defining it, so the definition, its basic properties, and its limits are established here.

Observation functors are ordinary functors from a category of systems to a category of observation records. The indistinguishability relation they induce is an equivalence relation whose classes are the record isomorphism classes of the functor.

Coarsening the observations identifies at least as many systems as the fine observations identify, and the inclusion of relations can be strict. The quantitative side of distinguishability splits along the classical and quantum divide: the Fisher metric is unique up to constant among metrics monotone under stochastic maps [8], while the quantum monotone metrics form a parametric family indexed by operator monotone functions [8]. An explicit two-dimensional computation shows the induced quadratic forms disagree on non-commuting directions, so distinguishability read from an unnamed monotone metric is not a function of the states alone.

A proved counterexample blocks the most natural equivalence criterion of all: equivalence of representation categories as plain tensor categories does not recover the underlying group, so any equivalence used to identify systems must state which layer of structure it preserves.

The separation between mathematical and physical equivalence runs through every section. Mathematical equivalence is isomorphism of representations. Physical equivalence is a relation to be chosen and defended. Observational equivalence is what an observation functor computes. The three notions must not be conflated. The series qualification on relational arguments binds throughout and is quoted in Section 7.

2 Equivalence notions

2.1 Systems and morphisms

Definition 2.1 (System record). A physical system representation is a tuple $P = (X, \mathcal{O}, \mathcal{D}, \mathcal{S}, \mathcal{C})$ comprising a state space or geometric structure X , an observable algebra or collection $\mathcal{O}$, dynamical data $\mathcal{D}$, symmetry and constraint data $\mathcal{S}$, and composition or interaction rules $\mathcal{C}$. This is the object schema used throughout the series. The five components need not be independent.

Morphisms between such records are the admissible relationships the program names: coarse-graining, exact equivalences, embeddings, dualities, and limiting procedures (the series morphism kinds). Their preservation laws are not fixed, so any statement below that quantifies over morphisms declares which morphism classes are in force. This paper uses three classes: functors of observation (Section 4), coarse-graining functors that forget microscopic detail while preserving macroscopic observables, and invertible equivalence morphisms. Set-theoretic size is handled once by convention: the class of admissible systems is assumed small enough for its quotients to be sets, or the arguments are read relative to a fixed universe.

¹The founding document of the YonedaAI Collaboration (2026, cited by section throughout this paper) and the YonedaAI physics research summary (2026-10-27) are program documents of the series in which this paper sits. Every mathematical statement used here is defined and stated explicitly in this text.

Relation	Notation	Identifies systems when	Status
Mathematical equivalence	$P \cong Q$	the tuples are isomorphic componentwise	established notion, strictly finer in general
Physical equivalence	$P \simeq Q$	the chosen relation holds	stipulative; the choice is open
Observational equivalence	$P \approx_{\text{Obs}} Q$	the observation records are isomorphic	defined here from an observation functor

Table 1: The three relations on system records. Only the first and third are computed from data; the second is a choice the classification problem depends on.

2.2 Equivalence relations on systems

Definition 2.2 (Mathematical equivalence). Two system records P and Q are mathematically equivalent, written $P \cong Q$, when their tuples are isomorphic componentwise in the ambient category of each component: a bijection of state spaces carrying observables, dynamics, symmetry data, and composition data across the identification. This is mathematical equivalence in the standard sense of isomorphism of representations.

Definition 2.3 (Physical equivalence). Physical equivalence, written $P \simeq Q$, is a carefully chosen equivalence relation on system records, used to form the quotient $\text{Ob}(\text{Phys})/\simeq$ in the classification problem. The relation is stipulated by choice; which choice is correct is open.

Definition 2.4 (Observational equivalence). Given an observation functor Obs from systems to a category of observation records (Theorem 4.1), two systems are observationally equivalent, written $P \approx_{\text{Obs}} Q$, when $\text{Obs}(P) \cong \text{Obs}(Q)$ in the category of observation records.

Definition 2.5 (Admissible candidate relation). Let $\simeq$ be a candidate relation claimed as physical equivalence, and let $\mathcal{A}$ be a family of admissible observation functors. The candidate is admissible only if

$$\cong \subseteq \simeq \subseteq \bigcap_{\text{Obs} \in \mathcal{A}} \approx_{\text{Obs}}.$$

The first inclusion states that an isomorphism of tuples is an identification no physical criterion can reject. The second states that physically equivalent systems are not separated by any admissible observation: if some admissible observation separates P and Q , the two systems report different outcomes to the same experiment, and declaring them physically equivalent contradicts the requirement that equivalent descriptions carry the same observable physics.

Table 1 summarizes the three relations, and two discipline points govern their use. Mathematical equivalence is strictly finer than physical equivalence in general: distinct mathematical descriptions can present the same observable physics. Observational equivalence is not a synonym for physical equivalence either; it is what one particular observation functor computes, and the choice of functor carries physical content (Section 8). The gap between the three relations is where the classification problem lives.

2.3 Preserved structure

Definition 2.6 (Equivalence witness). An equivalence witness for a proposed identification of P and Q is a record naming the equivalence notion in force (Theorem 2.2 or Theorem 2.3), the morphism kind realizing the identification, and the layers of structure the identification preserves.

The layers available are the symmetric monoidal structure, the monoidal (tensor) structure without symmetry, the braided structure, the $*$ -structure, the algebraic structure, the topological structure, complete positivity of channels, and the unitary freedom of minimal dilations. An identification that does not name its preserved layers is not admissible.

The requirement is substantive: an equivalence of representation categories that forgets the symmetric monoidal layer fails to recover basic data of the systems it identifies, by a proved counterexample rather than by a gap in technique (Section 3).

2.4 Morphism kinds at the observational level

The morphism kinds of the object schema act on observation records in characteristic ways, and stating the action is part of fixing the record category.

Definition 2.7 (Action of the morphism kinds). A coarse-graining is a functor c on the system category; the record of cP is computed by composing the observation functor with c . An embedding assigns to a subsystem its record and includes it in the record of the larger system. A duality carries records across presentations and must declare what it preserves (Section 3). A limiting procedure acts on records by a limit of records: the record of a thermodynamic or continuum limit is the limit of the records, when that limit exists in the record category.

Coarse-graining is the case Section 4.2 controls. Embeddings are controlled whenever the record of a composite system determines the records of its subsystems, which is a hypothesis on the record category rather than a theorem about observation. Dualities are controlled by the preserved-layer discipline of Section 3. Limiting procedures are the least settled case: a limit of records exists when the record category admits the relevant limits, and continuum limits in particular are not closed by formal means alone. Constructive quantum field theory treats continuum limits as essential, and nothing in this paper supplies the missing analysis. Any use of Theorem 2.7 must therefore say which morphism class is in force, exactly as the open status of the preservation laws requires.

2.5 Standing assumptions

The formal results of this paper rest on a short list of assumptions, and each is stated where it is used. The object schema of Theorem 2.1 is taken as the coordinate record of a representation, with the status of a stipulative schema rather than a settled definition of a physical system. The morphism classes in force are the three named at the end of Section 2. Statements quantified over admissible systems assume some fixed admissibility conditions; the condition list is left open in the series, so every such statement is conditional on the choice, and no result of this paper fixes the list.

The cited results are used as established mathematics with their hypotheses intact: the plain tensor counterexample (Theorem 3.1) with its construction cited rather than reproduced, the classical monotone-metric uniqueness (Theorem 5.1) on spaces of probability distributions, and the quantum parametric classification (Theorem 5.2) for operator monotone functions normalized at 1 and symmetric under $x \mapsto x^{-1}$ in Petz's sense. Nothing in this paper extends a cited result beyond its hypotheses, and the counterexample search of Section 6 keeps extension attempts visible.

3 Structure preservation and its failure

3.1 The preserved-layer requirement

Any equivalence used to identify systems in the classification quotient must state which layer of structure it preserves. The requirement exists because one natural criterion is false. The criterion

says: identify two systems when their representation categories are equivalent as tensor categories. The criterion is natural, since compositional structure is part of the object schema; it is also insufficient.

Theorem 3.1 (Plain tensor equivalence does not recover the group). *There exist finite groups G and H that are not isomorphic and whose complex representation categories are equivalent as plain tensor categories, that is, as monoidal categories without regard for the commutativity constraint. Consequently, an equivalence of representation categories that forgets the symmetric structure cannot recover the underlying group, and any classification that identifies systems by such an equivalence is false wherever the group data are observable superselection data of the systems.*

Proof, by reference. The construction of isocategorical groups and the proof appear in the original source [2], recorded in the series theorem registry as an obstruction result. The example pairs the affine symplectic group of a vector space over the field with two elements with a suitable affine pseudosymplectic group; the proofs run through triangular Hopf algebras [2]. The tensor-category equivalence is monoidal and does not respect the commutativity constraint [2]: the symmetric braiding is the layer the equivalence fails to preserve. This paper does not reprove the construction; it uses the statement as a constraint and searches for counterexamples to other criteria in Section 6. The series status of the statement is a complete proof in the mathematical literature, not a machine-checked result. $\square$

Corollary 3.2 (Preserved-layer discipline). *Every proposed physical equivalence $\simeq$ entering the quotient $\text{Ob}(\text{Phys})/\simeq$ must declare its preserved layers, and a declaration that omits the symmetric monoidal layer cannot justify identifying systems whose observable global symmetry data differ.*

Proof. If two systems were identified by an equivalence of representation categories as plain tensor categories, Theorem 3.1 exhibits pairs of groups for which such an identification misreports the systems as alike while their group data differ. When those data are observable (superselection sectors, global symmetry groups), the identification contradicts the observable physics. The declaration of preserved layers is the device that makes the failure visible before the quotient is formed. $\square$

The qualification in the corollary matters. Dualities of physical theories can identify systems whose presentation-level symmetry groups differ while the observable content agrees; the constraint forbids dropping the symmetric layer when the group data themselves are at stake, which is exactly the setting of Theorem 3.1.

Remark 3.3 (What the counterexample does not say). Theorem 3.1 removes the plain tensor layer as a recovery principle; it does not say tensor categories carry no group data. With the symmetric structure restored, representation categories with conjugates, direct sums, subobjects, and endomorphism ring of the tensor unit equal to the complex numbers are equivalent to finite-dimensional unitary representations of a compact supergroup, unique up to isomorphism [1]. The failure is specific: drop the symmetric layer and the group is no longer determined. The same point governs dualities below.

3.2 Duality and the recovered data

Dualities exchange structure between differently presented theories and are invertible only after passage to an appropriate opposite or dual category. Theorem 3.1 fixes how they must be recorded. A duality that carries one system to another is admissible as a physical identification only together with a statement of what it preserves; a duality that forgets the monoidal or symmetric layer of the compositional data is not an identification of physical systems, however strong the categorical

Preserved layer	What it determines	Status
Symmetric monoidal structure	the compact supergroup of the representation category under the stated hypotheses	reconstructible
Braided structure	more than the plain tensor layer, with extra soliton data in Galois settings	insufficient alone
Monoidal (tensor) structure without symmetry	composition of representations, not the group	provably insufficient
-structure and algebraic structure	the operator-algebraic presentation	needed for C^ -level statements
Complete positivity and unitary freedom	the channel and its minimal dilation up to unitary	normal-form level only

Table 2: Layers of structure an equivalence may preserve, and what each layer determines. An equivalence witness is checked against this list (Section 2).

equivalence. Braided settings need extra care again: the Galois theory of braided tensor categories shows that fixed-point-net representation categories require additional data beyond the braiding, recorded in the literature as soliton data [7].

Table 2 records the layers an equivalence may claim and what each layer determines. The third row is the operative one: the monoidal layer without symmetry is the one Theorem 3.1 forbids as a recovery principle, and the table is the checklist an equivalence witness is checked against.

The reconstruction pattern is the positive counterpart of the constraint. Gelfand duality, Stone duality, and the Gelfand-Naimark theorem recover named objects from named probes under named hypotheses [3, 4, 6]; Section 7 places that pattern against the observation formalism. None of these results establishes physical equivalence by itself; the series qualification on relational arguments governs every such step and is quoted in Section 7.

4 Observation functors

4.1 Definition and induced relation

Definition 4.1 (Observation functor). An observation functor is a pair $(\mathbf{Obs}, \mathbf{Obs})$ of a category $\mathbf{Obs}$ whose objects are observation records and whose morphisms are comparisons between records, together with a functor $\mathbf{Obs}: \mathbf{Phys}_{\text{adm}} \rightarrow \mathbf{Obs}$ that assigns to each admissible system its observation record and to each admissible morphism the induced comparison of records. An observation record packages the observational constraints or predictions a system presents to measurement: the family of accessible experiments, the outcomes each experiment reports, and the comparison of records under coarsening.

Theorem 4.1 is a proposal: the Information Geometry Correspondence research program names observation functors as its formal object without defining them. Two choices in the definition are deliberate. The domain is the admissible systems, not all of $\mathbf{Phys}$, because observation presupposes a system one can measure. The codomain is a category of records, not a set of numbers, because experiments compose and compare; a set-valued statistic would discard that structure.

The functorial framing differs from the decision-theoretic framing in which monotone metrics were first studied, and the difference is structural. The classical uniqueness of Section 5 originates in a decision-theoretic question about metrics on spaces of probability measures [8]: which metrics do not increase under stochastic transformations. That question presupposes the probability model and asks about its geometry. Theorem 4.1 moves one level up: it asks what data a system presents

to measurement before any metric is chosen. The two framings meet in Section 5, where the metrics are computed on the data Theorem 4.1 produces.

Proposition 4.2 (Induced relation). *For a fixed observation functor Obs , the relation $P \approx_{\text{Obs}} Q$ defined by $\text{Obs}(P) \cong \text{Obs}(Q)$ is an equivalence relation on admissible systems. The map $P \mapsto [\text{Obs}(P)]$ to record isomorphism classes is constant exactly on the $\approx_{\text{Obs}}$ classes, so the quotient by $\approx_{\text{Obs}}$ injects into the set of record isomorphism classes in the image of Obs , and every map of sets out of the systems that is constant on $\approx_{\text{Obs}}$ classes factors uniquely through that quotient.*

Proof. Isomorphism in Obs is an equivalence relation, and $\approx_{\text{Obs}}$ is its pullback along Obs : reflexivity, symmetry, and transitivity hold because they hold for isomorphism in Obs . By construction $P \approx_{\text{Obs}} Q$ holds exactly when $[\text{Obs}(P)] = [\text{Obs}(Q)]$, so the fibers of $P \mapsto [\text{Obs}(P)]$ (the record isomorphism classes, not the strict fibers of Obs on objects) are the $\approx_{\text{Obs}}$ classes, and the induced map on the quotient is injective on the image. Any set map constant on those fibers factors through the quotient by the universal property of quotient sets, and the factorization is unique because the quotient map is surjective. $\square$

The classes of $\approx_{\text{Obs}}$ are the fibers of the record-isomorphism-class map, sometimes called the essential fibers of Obs . Strict fibers, which fix one record object rather than its isomorphism class, are finer in general and are not the relation of Theorem 4.2.

Lemma 4.3 (Products of observations). *Let Obs_1 and Obs_2 be observation functors into record categories Obs_1 and Obs_2 . The assignment $P \mapsto (\text{Obs}_1(P), \text{Obs}_2(P))$ is an observation functor into the product category, and its induced relation is the intersection of the relations of the two factors: $P \approx Q$ for the product holds exactly when $P \approx_{\text{Obs}_1} Q$ and $P \approx_{\text{Obs}_2} Q$.*

Proof. The product of functors is a functor into the product category, and an isomorphism of pairs is exactly a pair of isomorphisms. The statement about relations is the definition of isomorphism in a product category. $\square$

Theorem 4.3 explains why adding experiments refines observational equivalence: each added observation functor intersects one more relation into the identification. Section 5 shows that in the quantum case even a single choice of distinguishability ruler is a choice of what counts as one measurement.

4.2 Coarse and fine observations

Each composition of observation with coarse-graining behaves in one direction only.

Definition 4.4 (Coarse and fine observations). Let Obs_f and Obs_c be observation functors into record categories Obs_f and Obs_c . The observation Obs_c is a coarsening of Obs_f when a functor $h: \text{Obs}_f \rightarrow \text{Obs}_c$ exists with $\text{Obs}_c = h \circ \text{Obs}_f$ on objects and morphisms. Separately, given a coarse-graining functor c on the system category (Theorem 2.7), the observation of the coarse-grained system is the composite functor $\text{Obs} \circ c$.

Proposition 4.5 (Coarse records separate less). *If Obs_c is a coarsening of Obs_f through h , then $P \approx_{\text{Obs}_f} Q$ implies $P \approx_{\text{Obs}_c} Q$, and the inclusion of relations can be strict: two systems with isomorphic fine records are identified by every coarsening, and two systems with the same coarse records can have non-isomorphic fine records.*

Proof. A comparison witnessing $\text{Obs}_f(P) \cong \text{Obs}_f(Q)$ maps under the functor h to a comparison witnessing $\text{Obs}_c(P) \cong \text{Obs}_c(Q)$, which proves the inclusion. For strictness, take $\Omega = \{1, 2, 3, 4\}$ with $p = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$ and $q = (\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8})$. Let Obs_f report the joint record of two experiments, the partition $\{\{1, 2\}, \{3, 4\}\}$ with pushforward $(\frac{3}{4}, \frac{1}{4})$ for both p and q , and the partition $\{\{1, 3\}, \{2, 4\}\}$ with pushforwards $(\frac{5}{8}, \frac{3}{8})$ and $(\frac{3}{8}, \frac{5}{8})$. Let h forget the second experiment, so Obs_c retains only the first. The fine records differ and the coarse records agree, and the coarsening h is exactly the forgetting map of Theorem 4.4. The computation is reproduced in the computational appendix (`observation_quotient.py`, function `coarse_vs_fine`). $\square$

Theorem 4.5 concerns coarsened observations of the same systems. It does not compare the records of P with the records of cP . A coarse-graining of systems is a different construction: fine-indistinguishable systems can have coarse-grained images that a fixed observation functor separates, because c need not respect the isomorphism of fine records. No statement of this paper identifies those two constructions.

4.3 Interaction and transitivity

Theorem 4.2 settles transitivity for functorially induced observational equivalence. The open question is whether indistinguishability generated by admissible experiments composes in interacting regimes. The series records this as unstudied, and the present paper does not close it. What can be said is conditional, and both regimes of the conditional statement are realistic.

Proposition 4.6 (Composed experiments). *Suppose the record category carries a product of records and the observation record of a composite system $A \otimes B$ is the product of the records of the factors. If $P \approx P'$ in the first factor relation and $Q \approx Q'$ in the second, then $P \otimes Q \approx P' \otimes Q'$. If instead the record of a composite contains a joint experiment whose outcome distribution is not determined by the factor records, the implication can fail.*

Proof. Under the stated hypothesis the record of a composite is a product of records, and a product of isomorphisms is an isomorphism, so indistinguishability of the factors is preserved. For the failure direction, take $\Omega = \{1, 2, 3, 4\}$ with distributions $p = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$ and $q = (\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8})$. Let the factor observation be the single experiment with outcome blocks $\{\{1, 2\}, \{3, 4\}\}$, under which $p \approx q$ (both push forward to $(\frac{3}{4}, \frac{1}{4})$). On the composite state space $\Omega \times \Omega$ the factor-marginal records of $p \times p$ and $p \times q$ coincide, yet the joint event $\omega_1 = \omega_2$ reports $\sum_i p_i^2 = \frac{11}{32}$ under $p \times p$ and $\sum_i p_i q_i = \frac{9}{32}$ under $p \times q$. The joint experiment separates two composites whose factor relations identify the factors. $\square$

Which of the two regimes holds is a property of the interaction data, not of the observation formalism. If the accessible experiments of an interacting composite are the products of the accessible experiments of the factors, composition is safe by Theorem 4.6. If the composition rules $\mathcal{C}$ of the object schema contribute experiments that neither factor admits alone, the implication has no reason to survive.

5 Information-geometric distinguishability

5.1 Classical uniqueness

Distinguishability between statistical models has a canonical classical measure. A Riemannian metric on a space of probability distributions is monotone under stochastic maps when no stochastic map increases the metric distance between distributions.

Theorem 5.1 (Chentsov). *Up to a positive constant multiple, the Fisher information metric is the unique Riemannian metric on spaces of probability distributions that is monotone under stochastic maps.*

Proof, by reference. The uniqueness statement is Chentsov's theorem, used here in the form stated and applied at [8]; the proof is a rigidity argument classifying monotone metrics on the classical simplex. This paper uses the statement as the classical anchor of distinguishability and does not reprove it. $\square$

Concretely, on the open simplex of positive probability vectors $p = (p_1, \dots, p_n)$, the Fisher form is

$$g_p^{\text{Fisher}}(a, a) = \sum_{i=1}^n \frac{a_i^2}{p_i}$$

for tangent vectors a with zero sum, and Theorem 5.1 says every monotone Riemannian metric on the simplex is a positive multiple of this form. The expression is the one that appears in Theorem 5.3 as the commutative restriction of the quantum family: on commuting data the quantum metrics all give the classical answer.

Classical statistical distance is therefore pinned down: any criterion that reads distinguishability from a monotone Riemannian metric on classical models must be a multiple of the Fisher metric. This is the one uniqueness result among the family of probability and measure-theoretic structures in the series taxonomy.

5.2 Quantum parametricity

In the quantum case the uniqueness fails. Monotone metrics on quantum state spaces are given by the Morozova-Chentsov construction indexed by operator monotone functions f on the positive reals, normalized by $f(1) = 1$ and self-reciprocal in Petz's sense, $f(x) = xf(x^{-1})$; the self-reciprocity is the symmetry condition that makes the quadratic form well defined on self-adjoint tangent directions, and distinct self-reciprocal f give distinct metrics.

Theorem 5.2 (Petz-Sudár classification). *Every quantum monotone metric arises from an operator monotone function f with $f(1) = 1$ and $f(x) = xf(x^{-1})$ through the Morozova-Chentsov form, and distinct such f give distinct metrics on self-adjoint directions. The family is parametric, not a finite list, and no member is distinguished among all monotone metrics.*

Proof, by reference. The classification of monotone metrics by operator monotone functions is stated and applied at [8], where the classical uniqueness and its quantum failure are contrasted. The self-reciprocity condition identifies each metric with a symmetric representative; without it, f and its transpose $f^\#(x) = xf(x^{-1})$ yield the same quadratic form on self-adjoint operators and would be double counted. The parametric nature of the family is the reason the classical uniqueness statement may not be extended to the quantum case. $\square$

Concretely, for a faithful density operator ρ with eigenvalues λ_i and a self-adjoint trace-free operator $A = \sum_{i,j} A_{ij}|i\rangle\langle j|$ in an eigenbasis of ρ , the quadratic form of the metric g_f is

$$g_f^\rho(A, A) = \sum_{i,j} \frac{|A_{ij}|^2}{\lambda_j f(\lambda_i/\lambda_j)}.$$

Proposition 5.3 (Commutative restriction). *For every operator monotone f with $f(1) = 1$, the quadratic form g_f^ρ restricted to operators A that are diagonal in an eigenbasis of ρ equals $\sum_i |A_{ii}|^2/\lambda_i$,*

the classical Fisher form on the commutative subalgebra generated by ρ . On commutative data the quantum monotone metrics therefore agree, and the parametric freedom of Theorem 5.2 is invisible.

Proof. For $i \neq j$ the summands of the quadratic form carry factors $|A_{ij}|^2 = 0$ when A is diagonal in the chosen basis. The surviving summands are $|A_{ii}|^2/(\lambda_i f(1)) = |A_{ii}|^2/\lambda_i$, since $f(1) = 1$. $\square$

5.3 Ruler dependence of the metric scale

The parametric freedom becomes visible on non-commuting data. Two self-reciprocal representatives suffice to see it: $f(x) = \frac{2x}{1+x}$ and $h(x) = \frac{1+x}{2}$, both operator monotone with value 1 at 1 and both satisfying $f(x) = xf(x^{-1})$.

Example 5.4 (Two monotone metrics disagree). Let $\rho = \text{diag}(3/4, 1/4)$ and $A = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then

$$g_f^\rho(A, A) = \frac{1}{\frac{1}{4} \cdot \frac{3}{2}} + \frac{1}{\frac{3}{4} \cdot \frac{1}{2}} = \frac{16}{3}, \quad g_h^\rho(A, A) = \frac{2}{\frac{3}{4} + \frac{1}{4}} + \frac{2}{\frac{1}{4} + \frac{3}{4}} = 4.$$

The two admissible monotone metrics give different quadratic forms at the same state and the same direction, by the factor 4/3.

Corollary 5.5 (Metric-level distinguishability is not canonical). *Any criterion that reads quantum distinguishability from a monotone metric must name the function f . Different admissible choices give different quadratic forms at the same states and directions, so the numerical scale of “the monotone-metric distance” is not a function of the states alone.*

Proof. Theorem 5.4 exhibits a single state and a single direction on which two admissible monotone metrics disagree, and Theorem 5.2 guarantees the family contains many more. A criterion built on an unnamed member of the family is therefore not a function of the systems alone. $\square$

The zero-distance relation is the same for every ruler: each g_f is positive definite on faithful states, so $d_f(\rho, \sigma) = 0$ exactly when $\rho = \sigma$, independently of f . The ruler-dependence lives in the numerical values of nonzero distances, not in the equality relation.

5.4 Distinguishability beyond metrics

The question behind the metric analysis is whether statistical distance yields an operational notion of distinguishability strong enough for classification. Of the available positions, the first two fail as routes to an equivalence relation.

A single monotone metric with a positive threshold ε defines a tolerance relation $\{(\rho, \sigma) : d_f(\rho, \sigma) \leq \varepsilon\}$. On a connected state space that relation is not transitive for any $\varepsilon > 0$, so by Theorem 4.2 it cannot be the relation of any observation functor, and it cannot enter the classification quotient. At $\varepsilon = 0$ the relation is state equality for every f , so the ruler plays no role. A family of rulers, one threshold per admissible f , has the same problem per ruler: each positive-threshold relation is non-transitive, and at zero threshold all rulers agree (Theorem 5.5 concerns the numerical scales, not the zero-distance relation). Theorem 4.3 applies to families of observation functors, not to tolerance relations, so it does not rescue the threshold construction.

Operational distinguishability defined by admissible experiments is the third position and the only one that produces equivalence relations through Theorem 4.2. It is also the least formalized: the admissible experiments of a general system class are not fixed. The position adopted here is the third as the construction of Section 4, with metrics as derived data computed on records rather than

Candidate criterion	Outcome	Evidence
Mathematical isomorphism as physical equivalence	fails to identify dual presentations	a space and its function algebra are non-isomorphic records presenting one object [3]; a system-level pair is not constructed here
Plain tensor-category equivalence	false as a recovery principle	Theorem 3.1, a proved counterexample
Monotone-metric distance without naming the metric	numerically not well defined in the quantum case	Theorem 5.4, computed and cross-checked
Coarse observation records as the whole record	identifies more than the fine record	Theorem 4.5 and its computational illustration

Table 3: Candidate criteria for physical equivalence and the result of the counterexample search.

as the definition of distinguishability. Metric thresholds remain legitimate as measures of nearness between records; they do not classify.

The parametricity result also fixes what a classifying invariant can be. A classifying invariant is a value that determines an object up to the chosen equivalence, computed from the object and nothing else. The Fisher metric has that status in the classical case: it is a function of the probability model alone, unique up to constant (Theorem 5.1). No quantum monotone metric has that status, because the value of the distance depends on the chosen function f in addition to the states (Theorem 5.5). The pair (state, ruler) can carry the invariant; the state alone cannot. Any classification table that lists a quantum distinguishability distance as an invariant of the states must therefore list the ruler as part of the invariant or drop the entry.

6 Counterexample search

The methodology of the series requires a counterexample search against each candidate criterion for physical equivalence. Table 3 records the candidates and the outcome.

The first row is only partially closed. The program’s founding document states that two mathematical descriptions may be physically equivalent without being mathematically isomorphic, and dual presentations are the motivating case. The table gives a mathematical illustration of dual presentations: a compact Hausdorff space and the commutative C^* -algebra of its continuous functions are distinct objects of their categories, isomorphic to no common record, yet Gelfand duality identifies them as presentations of one object [3]. A pair of system records inside one formal framework, proved non-isomorphic and declared physically equivalent, would close the row at the system level; the search did not produce one.

The second row is different in kind. Theorem 3.1 is a proved counterexample, and it falsifies the criterion. The third row is computational and analytic: Theorem 5.4 is verifiable by direct computation and is reproduced by the script named in the code appendix. The fourth row is the monotonicity of Theorem 4.5: a coarse record is computed from a fine one, so it cannot separate what the fine record identifies.

A further candidate deserves recording because it is tempting. One might define physical equivalence by equality of all observation records of a fixed maximal family of experiments, in the spirit of operational theories. This criterion is well defined once the family of experiments is fixed, and Theorem 4.2 applies to it directly. The difficulty is that “maximal family” is itself a choice: fix a strictly larger family and the relation becomes finer. No argument in the series selects a canonical

family for general systems, so the criterion restates the choice of $\simeq$ rather than resolving it.

6.1 A search protocol

The counterexample search above is repeatable for any future criterion. A candidate relation on system records comes with the morphism kind realizing its identifications, and the protocol screens it in three passes.

The screening pass reads the preserved layers off the morphism kind and checks them against Table 2; a witness naming the monoidal layer without symmetry fails on the strength of Theorem 3.1 immediately. The testing pass runs the candidate on four cases: a finite classical pair separated only by a second experiment, a qubit pair of monotone metrics disagreeing off the commutative slice, two non-isomorphic groups of order 8 with equal character tables, and a product of observations that separates what its factors hide. A criterion that misidentifies any case is false, and a criterion that passes all four is untested.

The band pass compares the candidate with Theorem 2.5: the candidate must be at least as coarse as isomorphism and no coarser than admissible observation.

A criterion that survives all three steps is not thereby correct. It is admissible so far, and the record of the search is part of the argument: which examples were run, which layers were named, which failure modes were ruled out. The series records no positive criterion for physical equivalence; the outcomes of Table 3 are the first run of the protocol.

7 Reconstruction patterns and representability

Observation functors are stipulated here rather than derived, because the two proved reconstruction patterns do not yet connect across general system classes.

The classic pattern recovers objects from probes. Gelfand duality reconstructs compact Hausdorff spaces from commutative C^* -algebras [3]; Stone duality reconstructs spectra from Boolean algebras [6]; the Gelfand-Naimark theorem embeds every C^* -algebra in an algebra of operators [4]. Each theorem is a precise statement with named hypotheses: an object is determined by how probes see it. The modern pattern represents higher-order operations. Categorical supermaps on a theory are concretely representable through channel-state duality whenever the theory has such a duality [5, 9], and the factorization theorems for supermaps rest on the Yoneda lemma and on representability [5].

Whether the two patterns connect into one principle is open. The series records the question explicitly: whether the categorical-supermap line unifies with the Gelfand, Stone, and Naimark pattern into a single account of objects determined by their relationships is not established, and this paper does not close it. The relevance to observation functors is a hypothesis rather than a result. If representability of higher-order operations extends from supermaps to observation records in general, then records could be defined as representable objects and the functorial account of Section 4 would be derived rather than stipulated. The hypothesis needs a channel-state-duality statement for the systems in question, and no such statement is on record for general system classes.

What can be used now is narrower. The reconstruction pattern fixes the shape of a legitimate claim about observation: an object recovered from named probes under named hypotheses is a theorem of reconstruction, not a classification and not an empirical adequacy claim. The series qualification applies to both patterns verbatim: “Yoneda does not, by itself, establish physical equivalence, empirical adequacy, or a classification theorem.” Section 4 defines observation functors as additional structure, and this section records the standing option of deriving them instead. The option is reasonable and unresolved.

8 The empirical layer

8.1 Instantiation of the layer

The three-layer architecture of the program separates mathematical classification, physical realization, and empirical selection. Layer three is a map from theories to observational constraints or predictions; it is not a claim that every model maps to a single observational object. This section instantiates layer three as an observation functor and marks where the instantiation stops.

Definition 8.1 (Empirical layer as observation functor). The empirical layer is the functor $\text{Obs}: \text{Phys}_{\text{adm}} \rightarrow \text{Obs}$ of Theorem 4.1, and the observational constraints or predictions of a theory are the contents of its observation record. The symbol $\mathcal{O}$ used for the codomain in the architecture collides with the observable-algebra symbol of the object schema; this paper writes **Obs** for the record category and keeps $\mathcal{O}$ for the observable algebra.

A mathematical structure can be consistent without describing our universe. Several mathematically distinct theories may produce identical predictions inside an experimentally accessible regime. A complete mathematical classification of representations would not by itself solve the empirical problem (founding document, section 7). The quotient $\text{Ob}(\text{Phys})/\simeq$ therefore cannot be formed by mathematical equivalence and then declared empirical; the relation must be argued for at the layer where the theories are compared.

Related work supplies a representability result in an adjacent direction. Categorical supermaps on a physical theory are concretely representable through channel-state duality whenever the theory has such a duality [5, 9]. That line shows observational structure of the higher-order kind is representable in the stated settings. It does not define observation records for general systems.

What the layer is mathematically remains open in the series, and the open items are specific. The composition of observation with coarsened observations is settled by functoriality (Theorem 4.5); the composition of empirical indistinguishability in interacting regimes is conditional (Theorem 4.6). The codomain **Obs** is proposed as a category of records here; the architecture allows observational constraints or predictions, and whether records are the right objects is a modeling choice the program has not fixed. The correct relation $\simeq$ is the central open problem of the whole series, and the results of this paper constrain it from three sides: it must be no finer than observation (Theorem 4.2), it must name its preserved layers (Theorem 3.2), and it must not rest on an unnamed monotone metric (Theorem 5.5).

8.2 Contents of the record category

Theorem 4.1 names the record category without fixing its objects beyond outcome data. The results of this paper force specific contents of that category, and each forced content names an open condition.

Experiments and outcomes are forced by Theorem 4.2: a record that carries no outcome data induces the universal relation, and every system is indistinguishable from every other. Outcome labels are fixed by calibration of a given experiment: isomorphism of records compares the outcome distributions of the same experiment on the same label set, and a relabeling of outcomes is an isomorphism of experiment presentations rather than of records of one calibrated experiment. Comparison morphisms between records are forced by Theorem 4.5: coarsening is a functor on records, so the record category must contain the morphisms that forgetting induces. Limits of records are forced only when limiting procedures are in force, and Section 2.4 records that case as the least settled.

What remains open is the treatment of interaction. The record of a composite system is not determined by the records of the factors unless the record category carries a monoidal structure mirroring the composition rules $\mathcal{C}$ of the object schema, and Theorem 4.6 shows both what holds under a product hypothesis and what can fail without it. The series leaves this open, and the record category proposed here inherits the open status.

8.3 Empirical selection as a quotient

Instantiated as an observation functor, the empirical layer computes a quotient: two theories land in the same class exactly when their records are isomorphic, which is the relation of Theorem 4.2. This is the part of empirical selection that observation performs, and its limits are the limits of the records. A class of the quotient contains every theory compatible with the recorded outcomes, and the founding document anticipates exactly this situation: several mathematically distinct theories may produce identical predictions inside an experimentally accessible regime (section 7).

Selection among the members of one class is not an observation problem in the sense of Theorem 4.1. It requires data the records do not carry: the internal structure of the theories, the behavior outside the accessible regime, or criteria of a different kind entirely. The three-layer architecture keeps this separation explicit: the empirical layer maps theories to observational constraints and predictions, and it is not a claim that every model maps to a single observational object. The quotient of Theorem 4.2 is the observational part of that map, nothing more.

9 Worked examples

The constructions of Sections 4 and 5 run on explicit data. The explicit data cover a finite classical observation functor with its quotient classes, the qubit computation of Theorem 5.4 in full generality, groups of order 8 whose character tables agree, and a product of observations.

9.1 A finite classical observation functor

A finite classical system is a probability vector p on a finite microstate set Ω , and its five-tuple of Theorem 2.1 is: X the simplex of distributions on Ω , $\mathcal{O}$ the experiment family, $\mathcal{D}$ trivial (static data), $\mathcal{S}$ trivial, and $\mathcal{C}$ the pointwise product of independent copies. An experiment is a map $e: \Omega \rightarrow \{1, \dots, k_e\}$, and its outcome distribution is the pushforward e_*p . Fix a family $E = \{e_1, e_2\}$ of experiments; the observation record of the system is the pair of outcome distributions $(e_{1*}p, e_{2*}p)$ with the outcome labels fixed by calibration, and two systems are observationally equivalent over E when these pairs agree. This is an observation functor into the category of finite outcome distributions with fixed labels, and Theorem 4.2 applies to it as stated.

Take $\Omega = \{1, 2, 3, 4\}$ and two distributions

$$p = (\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{8}), \quad q = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{8}, \tfrac{1}{8}).$$

Let e_1 partition Ω as $\{\{1, 2\}, \{3, 4\}\}$ and e_2 as $\{\{1, 3\}, \{2, 4\}\}$. Then $e_{1*}p = e_{1*}q = (\frac{3}{4}, \frac{1}{4})$, while $e_{2*}p = (\frac{5}{8}, \frac{3}{8})$ and $e_{2*}q = (\frac{3}{8}, \frac{5}{8})$. The coarse record consisting of e_1 alone identifies the two systems; the fine record over E separates them. This is the strictness claimed in Theorem 4.5, computed on four microstates.

The geometry of the classes is worth stating exactly, because the fine classes are not points. The class of p over the coarse record is the set of distributions with $p_1 + p_2 = \frac{3}{4}$, a two-dimensional slice of the simplex Δ_3 . The fine class of p over E is the line segment

$$p(t) = (\tfrac{1}{2} + t, \tfrac{1}{4} - t, \tfrac{1}{8} - t, \tfrac{1}{8} + t), \quad t \in [-\tfrac{1}{8}, \tfrac{1}{8}],$$

Direction (a, b)	$g_f^\rho(A, A)$	$g_h^\rho(A, A)$	Agreement
$(1, 0)$ (non-commuting)	$16/3$	4	no
$(0, 1)$ (commutative)	$16/3$	$16/3$	yes
$(1, 1)$ (mixed)	$32/3$	$28/3$	no
$(\frac{1}{2}, \frac{1}{2})$ (mixed)	$8/3$	$7/3$	no

Table 4: Quadratic forms of the two monotone metrics at $\rho = \text{diag}(3/4, 1/4)$ along four directions $A = a\sigma_x + b\sigma_z$. Agreement holds on commutative directions and fails elsewhere.

every point of which has $e_{1*}p(t) = (\frac{3}{4}, \frac{1}{4})$ and $e_{2*}p(t) = (\frac{5}{8}, \frac{3}{8})$. Adding the third partition $e_3 = \{\{1, 4\}, \{2, 3\}\}$ separates points of this segment, since $e_{3*}p(t) = (\frac{5}{8} + 2t, \frac{3}{8} - 2t)$ depends on t ; with $E' = \{e_1, e_2, e_3\}$ the record determines a distribution on Ω up to nothing, and the classes are single points.

9.2 The qubit computation

Theorem 5.4 computed one direction. The general two-dimensional case is closed, and the system tuple is: X the faithful qubit states, $\mathcal{O} = M_2(\mathbb{C})$, $\mathcal{D}$ trivial, $\mathcal{S}$ trivial, $\mathcal{C}$ the tensor product. Let $\rho = \text{diag}(\frac{3}{4}, \frac{1}{4})$ and $A = a\sigma_x + b\sigma_z$ with real a, b , so $A = \begin{pmatrix} b & a \\ a & -b \end{pmatrix}$. For $f(x) = \frac{2x}{1+x}$ the quadratic form of Section 5 is

$$g_f^\rho(A, A) = \frac{16}{3}(a^2 + b^2),$$

and for $h(x) = (1+x)/2$ it is

$$g_h^\rho(A, A) = \frac{16}{3}b^2 + 4a^2.$$

The two metrics differ only on off-diagonal directions. The diagonal summands contribute $\frac{4}{3}b^2 + 4b^2 = \frac{16}{3}b^2$ in both cases. For $f(x) = \frac{2x}{1+x}$ each of the two off-diagonal summands equals $\frac{a^2}{\lambda_2 f(\lambda_1/\lambda_2)} = \frac{8}{3}a^2$, so the off-diagonal total is $\frac{16}{3}a^2$; for $h(x) = (1+x)/2$ the two summands are $2a^2$ each, for a total of $4a^2$. The metrics coincide when $a = 0$ and differ by the factor $4/3$ when $b = 0$.

Table 4 records the values on four directions. The pattern is the content of Theorem 5.3 and Theorem 5.5 together: the rulers coincide on commutative data and disagree everywhere else, so a distinguishability criterion that reads a monotone metric must name its function f .

9.3 Groups of order 8

Let D be the dihedral group of order 8 with presentation $\langle r, s \mid r^4 = s^2 = 1, srs = r^{-1} \rangle$ and Q the quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}$. Both have five conjugacy classes, class sizes $1, 1, 2, 2, 2$, abelianization $C_2 \times C_2$, and hence four one-dimensional characters; the remaining irreducible character has degree two and values $(2, -2, 0, 0, 0)$ on the classes in the order (identity, central involution, three classes of size two). The two character tables are therefore the same matrix of numbers once characters and the three size-2 classes are ordered; Table 5 displays one such ordering. The groups are not isomorphic: D has five elements of order 2 and two of order 4, while Q has one element of order 2 and six of order 4.

The character table does not separate D and Q . The computation in `invariant_coarseness.py` derives both tables from the multiplication tables, by enumerating conjugacy classes, extracting the one-dimensional characters from the abelianization, and completing the table against the regular character; it then compares the tables and the element-order data. This finite statement is proved by the computation. It is a shadow of Theorem 3.1, not the theorem itself: the tensor-category

Character	identity	central involution	size-2 class	size-2 class	size-2 class
χ_1	1	1	1	1	1
χ_2	1	1	1	-1	-1
χ_3	1	1	-1	1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

Table 5: The character table shared by the dihedral group of order 8 and the quaternion group of order 8, classes ordered as described in the text. Both groups realize this table.

counterexample lives at the level of monoidal equivalences that fail to respect the commutativity constraint [2], and equality of character tables is a strictly weaker coincidence.

9.4 Products of observations

Theorem 4.3 says that adding an experiment intersects a relation into the identification. The four-microstate data of Section 9 shows the mechanism on three systems. Take $p = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$, $q = (\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8})$, and $r = (\frac{1}{8}, \frac{1}{8}, \frac{1}{2}, \frac{1}{4})$. Over the single experiment e_1 the records of p and q coincide at $(\frac{3}{4}, \frac{1}{4})$ and the record of r is $(\frac{1}{4}, \frac{3}{4})$, so the relation identifies p with q and separates r . Over the single experiment e_2 the records of p and r coincide at $(\frac{5}{8}, \frac{3}{8})$ and the record of q is $(\frac{3}{8}, \frac{5}{8})$, so the relation identifies p with r and separates q . Over the product of the two experiments the three records are pairwise distinct, and the relation is the intersection of the two single-experiment relations: only the identity remains on the three systems. Each experiment hides one coincidence that the other exposes, and no coincidence survives the product. This example concerns one system under two experiments; the composite-system failure of Theorem 4.6 is a different construction, exhibited in its proof.

10 Computational evidence

The computational investigations accompany the analytic arguments as evidence, not proof: the claims they support stand or fall on the proofs in Sections 4 and 5, and the computations check the arithmetic and exhibit the constructions on explicit finite data. The scripts live in `src/vol-10-observational-equivalence/` and are listed in the code evidence appendix (Table 6).

The first script builds a finite toy category of classical systems with observation maps and checks that the induced indistinguishability is an equivalence relation, that the quotient classes match the fibers of the record map, and that coarse observations separate strictly less than fine ones. This is the finite shadow of Theorems 4.2 and 4.5.

The second script evaluates the Morozova-Chentsov quadratic form of Theorem 5.4 for the two functions $f(x) = \frac{2x}{1+x}$ and $h(x) = (1+x)/2$ at the state $\rho = \text{diag}(3/4, 1/4)$ in the direction $A = \sigma_x$, and reproduces the values $16/3$ and 4 . It also evaluates the four directions of Table 4 and checks the closed forms of Section 9.

The third script computes character tables of the dihedral group of order 8 and the quaternion group of order 8 from their multiplication tables, by brute-force enumeration of conjugacy classes, one-dimensional characters from the abelianization, and completion of the table by orthogonality against the regular character. The two groups are not isomorphic (their element orders differ), and their character tables coincide as tables of numbers. The computation illustrates the failure of coarse representation data to separate groups. It is not the counterexample of Theorem 3.1: that

Claim supported	File	Function	What it demonstrates
Induced indistinguishability is an equivalence relation with record-class fibers	<code>observation_quotient.py</code>	<code>check_axioms</code>	axiom checks on a finite toy category
Coarse observations separate strictly less than fine ones	<code>observation_quotient.py</code>	<code>coarse_vs_fine</code>	an explicit pair identified by coarse records only
Monotone quadratic forms depend on f off the commutative subalgebra	<code>monotone_metric_contrast.py</code>	<code>mc_quadratic_form</code>	reproduces 16/3 and 4 of Theorem 5.4
Character tables do not separate the two groups of order 8	<code>invariant_coarseness.py</code>	<code>compare_character_tables</code>	equal tables, distinct element-order data

Table 6: Code evidence appendix. Each row names a claim, the file and function that demonstrates it, and what the demonstration shows. The computational evidence is not proof.

counterexample lives at the level of monoidal categories and their symmetric structure, and the script does not construct it.

The Haskell module `Instances.hs` in the same directory declares instances of the shared family classes of the series taxonomy for the systems this paper uses: observation functors over classical and quantum records, the equivalence classes of Theorem 4.2, and the witness type that records preserved layers in the sense of Theorem 3.2. The module is an executable model of the definitions; it proves nothing and is not verified by any proof assistant.

11 Constraints on the classification quotient

The classification problem is the quotient of admissible systems by $\simeq$. The results of this paper do not determine $\simeq$, and Theorem 2.5 bounds where it can sit.

Theorem 2.5 is a constraint, not a construction, and both of its failure modes are live. A relation too close to $\cong$ cannot identify dual presentations, and the founding document records such presentations as the reason $\simeq$ is a choice at all. A family $\mathcal{A}$ that is too small over-identifies: the fewer observation functors are admitted, the larger the intersection, and a system class with one weak experiment identifies systems no working physicist would call the same. Enlarging $\mathcal{A}$ shrinks the intersection toward $\cong$ (Theorem 4.3 is the mechanism), so an oversized family can empty the band for the dual presentations that $\simeq$ is supposed to identify. The band is usable only for families of observations that match the physics in question.

The counterexample search of Section 6 adds a third constraint: a candidate that realizes its identifications by plain tensor equivalence violates Theorem 3.1 wherever the group data are observable, and a candidate that reads numerical distances from an unnamed monotone metric is not well defined in the quantum case (Theorem 5.5).

The preserved-layer discipline is the one constructive tool the paper offers. A candidate $\simeq$ becomes checkable when its proponents file an equivalence witness (Theorem 2.6) for each identification: the notion in force, the morphism kind, the preserved layers. Isocategorical groups then serve as the standing test: a witness whose preserved layers omit the symmetric monoidal structure and whose systems carry different observable group data must be rejected before further analysis. The series records no positive reconstruction criterion for “the same physical system”; the search for one is open, and the constraints above are the admissibility conditions any such criterion must meet.

The relation to the three-layer architecture is immediate. The band of Theorem 2.5 sits at the empirical layer: the first inclusion is mathematical, the second is empirical, and neither follows from

the other. A complete classification of the mathematical representations (layer one) would leave the second inclusion unproved, and a proof of the second inclusion would leave the quotient of layer one uncomputed.

11.1 What would settle the choice

The choice of $\simeq$ moves from open to fixed only through concrete tasks, each stated so that it can fail. The first task is to fix the record category for a named class of systems: the admissible experiments, the isomorphisms of records, and the composition of records under interaction. Success means a construction to which Theorem 4.2 applies outside the toy model of Section 9; failure means the class admits no well-formed observation functor. The second task is to determine whether bounded experimental indistinguishability composes in an interacting regime, the question Section 4 leaves open through Theorem 4.6. Success means a theorem connecting the functorial relation of Section 4 with an operational relation between records; failure means a counterexample in which two systems are indistinguishable by every individual experiment and distinguishable by a joint one. The third task is to state a positive criterion for physical equivalence and run it through the protocol of Section 6.1. Success means a criterion that survives the protocol and is argued for at the empirical layer; failure means another row of Table 3 closes negatively, which is itself a result the program accepts.

Until the three tasks are run, the constraints of this section are the result: a band the correct relation must lie in, a discipline for declaring what identifications preserve, and a standing counterexample to the most natural criterion of all.

12 Limitations

The observations of this paper are proposals constrained by proved results; they are not a completed formalization of the empirical layer. The domain and codomain of Theorem 4.1 fix the shape of an observation functor, and the sources do not fix the contents: which experiments are admissible for a given system, how outcome families compose under interaction, and what structure the record category must carry are open. Until they are fixed, Theorem 4.2 is a statement about a proposed construction, not about physics.

The counterexample search found one proved falsification (Theorem 3.1) and one computed failure of canonicity (Theorem 5.4); it did not produce a concrete non-isomorphic pair of system records to close the dual-presentation row of Table 3 at the system level. The search is bounded by the sources reviewed (cutoff 2026-10-29), and the absence of such a pair in this paper is not evidence that the literature lacks one.

The metric results carry their own boundary. Theorems 5.1 and 5.2 are used as established mathematics with their hypotheses intact: monotone under stochastic maps in the classical case, operator monotone functions normalized at 1 and self-reciprocal in the quantum case. Nothing here extends classical uniqueness to the quantum case, and Theorem 5.5 is the reason not to. The explicit computation is two-dimensional; the disagreement of the metrics is general (Theorem 5.2) but only the two-dimensional instance is exhibited here.

The observational equivalence defined here is functorial indistinguishability. Bounded experimental indistinguishability in interacting regimes is not shown to coincide with it (Theorem 4.6), and the empirical layer of the three-layer architecture keeps its full open status wherever the sources keep it.

13 Conclusion

Observational equivalence can be defined with ordinary categorical tools: an observation functor induces an equivalence relation whose classes are its record isomorphism classes, and coarsening the observations identifies at least what the fine observations identify. Distinguishability has a canonical classical ruler and no canonical quantum one: the Fisher metric is unique among monotone metrics on classical models, while quantum monotone metrics form a parametric family that agrees only on commutative data and disagrees on explicit non-commuting examples. Physical equivalence remains a choice, constrained from three sides by the results of this paper: it cannot be finer than observation, it must name the layers of structure it preserves, and it cannot rest on an unnamed metric. The most natural categorical criterion for it is false by a proved counterexample: plain tensor equivalence does not recover the group.

What follows is a research program rather than a theorem. Fix the admissible experiments of a system class, build the record category for that class, and ask whether the induced quotient is the one the classification problem needs. Until then, the quotient $\text{Ob}(\text{Phys})/\simeq$ carries the equivalence relation as the open choice the quotient turns on, and papers in this series must state which relation they use.

Notation

Symbol	Meaning
P	system record $P = (X, \mathcal{O}, \mathcal{D}, \mathcal{S}, \mathcal{C})$: state space, observables, dynamics, symmetries, composition rules (Theorem 2.1)
$P \cong Q$	mathematical equivalence: isomorphism of tuples (Theorem 2.2)
$P \simeq Q$	physical equivalence: the chosen relation of the classification quotient (Theorem 2.3)
$\text{Obs}, \mathbf{Obs}$	observation functor and its category of records (Theorem 4.1)
$P \approx_{\text{Obs}} Q$	observational equivalence induced by Obs (Theorem 2.4)
$\approx_{\text{Obs}_c}$	relation of a coarsened observation (Theorem 4.4)
$g_f^p(A, A)$	Morozova-Chentsov quadratic form of the monotone metric g_f (Section 5)
f, h	self-reciprocal operator monotone functions normalized at 1: $f(x) = \frac{2x}{1+x}$ and $h(x) = \frac{1+x}{2}$
$\mathcal{A}$	family of admissible observation functors (Theorem 2.5)

Table 7: Notation used throughout the paper.

References

- [1] S. Doplicher, J. Roberts, *A new duality theory for compact groups*, *Inventiones Mathematicae* 98 (1989) 157–218.
- [2] P. Etingof, S. Gelaki, *Isocategorical groups*, arXiv:math/0007196, 2000-07-31.
- [3] I. Gelfand, *Normierte Ringe*, *Recueil Mathematique* 9(51):1 (1941) 3–24.
- [4] I. Gelfand, M. Neumark, *On the imbedding of normed rings into the ring of operators in Hilbert space*, *Recueil Mathematique* 12(54):2 (1943) 197–217.

- [5] J. Hefford, M. Wilson, *A Profunctorial Semantics for Quantum Supermaps*, arXiv:2402.02997, 2024-02-05.
- [6] P. Johnstone, *Stone Spaces*, Cambridge Studies in Advanced Mathematics 3, Cambridge University Press, 1982.
- [7] M. Müger, *Conformal Field Theory and Doplicher-Roberts Reconstruction*, arXiv:math-ph/0008027, 2000-08-21.
- [8] D. Petz, Cs. Sudár, *Extending the Fisher metric to density matrices*, arXiv:quant-ph/0102132, 2001-02-27.
- [9] M. Wilson, J. Hefford, T. Hoffreumon, *Supermaps on generalised theories*, arXiv:2602.23865, 2026-02-27.