Distinguishability Geometry: Projection, Embedding, Revision, and Transport

Flyxion

Independent Researcher

June 2026

Abstract

We propose *distinguishability geometry* as a unified framework for understanding how representational systems gain, lose, and transfer the capacity to express distinctions. The foundational primitive is the *ontological triple* $(X, \sim, T)$: a set X, an observational indistinguishability relation $\sim$, and a generating ontology T. Every such triple carries two intrinsic invariants: the *coding deficit* $\delta_T(x) = L_T(x) - L^*(x)$, measuring excess description length, and the *distinguishability deficit* $\Delta_T(x) = D^*(x) - D_T(x)$, measuring inaccessible distinction capacity. The *Deficit Proxy Theorem* establishes $\Delta_T(x) \leq C\,\delta_T(x)$, making the coding deficit a computable surrogate for the geometric one; full equivalence is conjectured but not proved.

We identify four primitive operations arranged in a two-level hierarchy. At the *distinction level*, projection and embedding coarsen and refine $\sim$ within a fixed space. At the *ontology level*, revision restructures the generating ontology and transport maps distinguishability structure across spaces. The *Representational Reduction Proposition* establishes these as exhaustive qualitative fiber effects; the *Classification Conjecture* asserts every representational transformation factors into their finite composition.

For each operation we derive a characteristic theorem: a Projection Deficit Bound in entropic form, an Embedding Refinement theorem, a Revision Monotonicity theorem with asymptotic corollary, and a Transport Stability theorem giving a Lipschitz condition for analogy. A *Distortion–Deficit Relationship* connects the two error measures. A concrete finite example illustrates all four operations. We close with a categorical formulation in which the hierarchy corresponds to morphism types and the deficit is a functor from the category **Dist** of ontological triples.

Contents

1. Introduction

Three independent research programmes have converged on a common mathematical object. The first, from process-primary computation, establishes that observable state is a projection of a richer history space and that distinct histories collapse to the same state. The second, from kernel statistics, demonstrates that distinct distributions become mutually singular after covariance embedding, amplifying separation without creating information. The third, from archive-based compression, introduces *reachable complexity*: the shortest description expressible from the archive's current template hierarchy, which may be strictly longer than the Kolmogorov-optimal description.

Each programme arrives at the same hidden structure from a different direction: the *geometry of distinguishability*. The first studies distinctions *destroyed* by representation. The second studies distinctions *amplified*. The third studies distinctions that are *inaccessible* within the current frame.

The present paper makes this object explicit. Our main contributions are:

(1) The *ontological triple* $(X, \sim, T)$ as the correct object of study, fixing the under-specification of earlier formulations that left the generating ontology implicit.

(2) Two intrinsic invariants—coding deficit and distinguishability deficit—and a *Deficit Proxy Theorem* relating them.

(3) A two-level hierarchy of four operations with characteristic theorems for each, including a corrected Embedding Refinement proof, an entropic Projection Deficit Bound, and a Lipschitz Transport Stability theorem.

(4) A *Distortion–Deficit Relationship* connecting the two error measures that previously lived in separate worlds.

(5) A *Classification Conjecture* precisely stated as a factorization claim, together with a worked finite example illustrating all four operations.

(6) A categorical formulation of **Dist** with well-defined objects and morphism types.

The paper is organized as: spaces and invariants (§2), classification conjecture (§3), four operations with theorems (§4), asymmetries and meta-theorem (§5), deficit as universal invariant (§6), and categorical formulation (§7).

2. Distinguishability Spaces and Their Invariants

2.1. The Ontological Triple

Earlier formulations defined the foundational object as a pair $(X, \sim)$. This is insufficient: the indistinguishability relation $\sim$ on X does not by itself determine which descriptions are reachable, and hence does not determine the deficit. The correct object is a triple.

Definition 2.1 (Ontological triple)**.** An *ontological triple* is $(X, \sim, T)$ where X is a set, $\sim$ is an equivalence relation on X (observational indistinguishability), and T is an ontology: a description language or template hierarchy determining which descriptions of elements of X are reachable. The equivalence classes $[x]_\sim$ are called *fibers*.

The same pair $(X, \sim)$ can appear in multiple triples with different ontologies T, T', inducing the same fiber structure but different deficit values. This is not a defect; it reflects the genuine dependence of representational cost on the description language.

Proposition 2.2 (Fiber Monotonicity)**.** *Let* $(X, \sim_1, T)$ *and* $(X, \sim_2, T)$ *share the same set and ontology with* $\sim_1 \subseteq \sim_2$ *(*$\sim_2$ *coarser). Then:*

$$[x]_{\sim_1} \subseteq [x]_{\sim_2}, \qquad |[x]_{\sim_1}| \leq |[x]_{\sim_2}|.$$

Proof. If $y \in [x]_{\sim_1}$ then $x \sim_1 y$; since $\sim_1 \subseteq \sim_2$, $x \sim_2 y$, so $y \in [x]_{\sim_2}$. Cardinality follows. □

Corollary 2.3. *Coarsening a distinguishability relation never increases the number of distinguishable classes. Projection always reduces expressible distinctions.*

2.2. The Distinguishability Principle

Principle 2.4 (Distinguishability Principle). A representational system is characterized not by what objects it contains but by what distinctions it can express.

Every ontological triple possesses a deficit because no finitary ontology can express all distinctions present in the world.

2.3. Two Deficit Notions

Definition 2.5 (Coding deficit)**.** Let $L_T(x)$ be the shortest description of x reachable within T, and $L^*(x) = K(x)$ the Kolmogorov complexity of x. The *coding deficit* is:

$$\delta_T(x) \;=\; L_T(x) - L^*(x).$$

Definition 2.6 (Distinguishability deficit)**.** Let $D_T(x)$ be the distinguishability capacity of ontology T at x: the number of objects separable from x by descriptions reachable in T. Let $D^*(x)$ be the maximum capacity attainable by any ontology. The *distinguishability deficit* is:

$$\Delta_T(x) \;=\; D^*(x) - D_T(x).$$

Theorem 2.7 (Deficit Non-Negativity)**.** *For every ontological triple* $(X, \sim, T)$ *and every* $x \in X$*:*

$$\delta_T(x) \geq 0, \qquad \Delta_T(x) \geq 0.$$

Both vanish iff T *is locally optimal for* x*.*

Proof. $L^*(x)$ is the minimum over all ontologies, so $L^*(x) \leq L_T(x)$, giving $\delta_T(x) \geq 0$. Similarly $D^*(x)$ is the maximum, so $D_T(x) \leq D^*(x)$, giving $\Delta_T(x) \geq 0$. Both equalities hold iff T achieves the respective optimum. □

Theorem 2.8 (Deficit Proxy)**.** *Under the assumption that each expressible distinction requires at least one description bit (*faithful encoding*), the coding deficit bounds the distinguishability deficit:*

$$\Delta_T(x) \;\leq\; C\,\delta_T(x)$$

for a system-dependent constant $C > 0$*. In particular:*

$$\delta_T(x) = 0 \implies \Delta_T(x) = 0.$$

The converse direction, $\Delta_T(x) = 0 \Rightarrow \delta_T(x) = 0$*, requires the additional assumption that optimal coding uses all available distinctions, and is left as a conjecture.*

Proof. Under faithful encoding, each inaccessible distinction in $\Delta_T(x)$ requires at least one additional bit in T's description of x to flag the corresponding fiber collapse. Hence $\Delta_T(x) \leq \delta_T(x)$, and the general bound $\Delta_T(x) \leq C\,\delta_T(x)$ follows with $C \geq 1$ scaling with the maximum description-length cost per inaccessible distinction.

For the forward implication: if $\delta_T(x) = 0$, then T achieves optimal description length for x. Under faithful encoding, this means every bit-costly distinction is expressible in T, so no distinction is inaccessible: $\Delta_T(x) = 0$. □

Conjecture 2.9 (Full Deficit Correspondence)**.** Under faithful encoding, $\Delta_T(x) = 0 \iff \delta_T(x) = 0$.

Remark 2.10. The Deficit Proxy Theorem is sufficient for all subsequent arguments in this paper. The coding deficit $\delta_T(x)$ is a computable surrogate for the geometric quantity $\Delta_T(x)$, and the two share the same zero set under the proxy direction. Full equivalence remains an open problem.

2.4. A Concrete Illustration

Before stating the classification conjecture, we ground the framework in a finite example that will be extended throughout the paper.

Example 2.11 (Four operations on a four-element set)**.** Let $X = \{a, b, c, d\}$ with ontology T_0 admitting a full distinguishing description for each element. Consider the following relations and ontologies:

Initial space. $\sim_0 = \big\{\{a\}, \{b\}, \{c\}, \{d\}\big\}$ (discrete: all elements distinguishable). Here $\delta_{T_0}(x) = 0$ for all x.

Projection. The map $\sigma : X \to \{p, q, r\}$ with $\sigma(a) = \sigma(b) = p$, $\sigma(c) = q$, $\sigma(d) = r$ induces $\sim_1 = \big\{\{a, b\}, \{c\}, \{d\}\big\}$. The fiber $F_p = \{a, b\}$ has size 2, so the projection deficit bound gives $\Delta\delta(\sigma) \geq \log 2 = 1$ bit.

Embedding. From $(X, \sim_0)$, map $\phi : X \to X \times \{0, 1\}$ by $\phi(x) = (x, f(x))$ where $f(a) = f(c) = 0$, $f(b) = f(d) = 1$. This refines $\sim_0$ by separating elements according to a new attribute: $\phi(a) \not\sim' \phi(b)$ even when $a \sim_0 b$ under a coarser version of $\sim_0$. A faithful embedding here preserves all existing distinctions and adds new ones.

Revision. Starting from $(X, \sim_1, T_0)$ with $\sim_1 = \big\{\{a, b\}, \{c\}, \{d\}\big\}$, suppose a new ontology $T_1 \supset T_0$ adds a template that distinguishes a from b directly. Then $\sim_{T_1} = \sim_0$ is recovered and $\delta_{T_1}(a) = \delta_{T_1}(b) = 0$: a productive revision has decreased the deficit.

Transport. Let $Y = \{w, x, y, z\}$ with $\sim_Y = \big\{\{w, x\}, \{y\}, \{z\}\big\}$. The map $\tau : X \to Y$ given by $\tau(a) = w$, $\tau(b) = x$, $\tau(c) = y$, $\tau(d) = z$ is a transport with zero distortion: it preserves the fiber structure of $\sim_1$ exactly ($a \sim_1 b \iff \tau(a) \sim_Y \tau(b)$).

Hybrid operation. Note that the transformation $\sim_0 \to \sim_3$ where $\sim_3 = \big\{\{a, b\}, \{c, d\}\big\}$ is *not* a pure projection (it coarsens in two places simultaneously). By the Classification Conjecture (§3), it should factor as a composition: e.g. first project $\{c, d\}$ to a single class, then project $\{a, b\}$. This factorization is not unique, but the conjecture asserts it always exists.

3. The Classification Conjecture

Before developing the four operations in detail, we state the conjecture that organizes the entire framework. The preceding example and the Representational Reduction Proposition (stated as Proposition 4.1 in §4) motivate it; the following sections constitute the evidence for it.

Conjecture 3.1 (Classification Conjecture)**.** Every representational transformation factors into a finite composition of projection, embedding, revision, and transport maps.

Three clarifications are essential.

Factorization, not identity. The transformation itself need not be one of the four primitive operations. A hybrid transformation that simultaneously coarsens some fibers and refines others factors as a composition—for instance, a projection followed by an embedding. This is precisely the sense in which the conjecture resembles factorization theorems in algebra and topology.

Scope. The conjecture concerns representational transformations: maps between ontological triples $(X, \sim, T)$ that may alter any or all of X, $\sim$, and T. Pure set-theoretic maps that ignore the relational and ontological structure are outside its scope.

Evidence. The characteristic theorems of §4 can be read as partial evidence: they show that the four operations, together, are capable of expressing the full range of qualitative fiber behaviors. Whether every transformation *factors* into these behaviors—whether no genuinely new operation can arise from their combination—is what the conjecture claims and what remains to be proved.

4. The Four Operations

4.1. The Two-Level Structure

The four operations divide into two levels according to what they modify:

$$\text{Objects} \longleftrightarrow \text{Distinctions} \longleftrightarrow \text{Ontologies}.$$

Projection and embedding operate at the *distinction level*: they modify $\sim$ within a fixed space, moving horizontally between finer and coarser relations on the same X. Revision and transport operate at the *ontology level*: revision changes T (moving vertically in the diagram) and transport connects two distinct triples (moving between nodes).

Proposition 4.1 (Representational Reduction)**.** *Every representational transformation induces a transformation on distinguishability structure. The qualitatively distinct fiber effects are:*

- ***Projection****: merges fibers (coarsens* $\sim$*).*
- ***Embedding****: separates fibers (refines* $\sim$*).*
- ***Revision****: restructures the ontology generating* $\sim$*.*
- ***Transport****: maps fiber structure between distinct triples.*

These four effects are mutually exclusive as pure operations and jointly exhaust the qualitative possibilities.

Level	**Operation**	**Effect on** $\sim$	$\Delta\delta$	**Failure mode**
Distinction	Projection	Coarsens $\sim$	≥ 0	Information loss
Distinction	Embedding	Refines $\sim$	≤ 0	Spurious distinctions
Ontology	Revision	Restructures T	< 0 (productive)	Lock-in
Ontology	Transport	Maps $\sim$ across spaces	bounded	Fiber distortion

Table 1: The four operations, their level, deficit change $\Delta\delta(f)$, and canonical failure modes. Note: $\epsilon(f) \geq 0$ always; $\Delta\delta(f)$ is signed.

4.2. Formal Separation: Distortion vs. Deficit Change

Definition 4.2 (Distortion and deficit change)**.** For a map f between ontological triples:

- $\epsilon(f) \geq 0$: the *distortion* of f, measuring how much f fails to be a fiber-isomorphism. Always non-negative.
- $\Delta\delta(f) = \delta_{T'}(f(x)) - \delta_T(x)$: the *deficit change* induced by f. Signed: positive means deficit increases, negative means it decreases.

These quantities are related. The following theorem provides the connection that was previously missing.

Theorem 4.3 (Distortion–Deficit Relationship)**.** *Let $f : (X, \sim_X, T_X) \to (Y, \sim_Y, T_Y)$ be a representational transformation with distortion $\epsilon(f)$. Then under faithful encoding:*

$$|\Delta\delta(f)| \;\leq\; K\,\epsilon(f)$$

for a constant $K > 0$ depending on the description complexity of individual fibers in T_X and T_Y.

Proof. Each unit of distortion $\epsilon(f)$ corresponds to a fiber violation: either a pair that was equivalent in X becoming inequivalent in Y (distinction inflation) or a pair inequivalent in X becoming equivalent in Y (distinction collapse). Under faithful encoding, each such violation costs at most K bits of description length: collapsing a fiber saves at most $\log |F|$ bits (gained by the projection) and inflating creates at most $\log |F'|$ bits (cost of the new separation). In both cases, the change in deficit is bounded by the description complexity of the affected fibers, which is controlled by $K \cdot \epsilon(f)$ where K is the maximum per-violation description cost. □

Remark 4.4. Theorem 4.3 makes distortion and deficit change into *exchange rates* between geometry and information. A transformation with small distortion cannot change the deficit by much; a large deficit change must be accompanied by substantial distortion. This connects the geometric and coding-theoretic aspects of the framework into a single inequality.

4.3. Projection

Definition 4.5 (Projection)**.** A *projection* on $(X, \sim, T)$ is a surjective map $\sigma : X \to S$ inducing a coarsening $\sim'$ of $\sim$: $\sigma(x) = \sigma(y) \implies x \sim' y$. Fibers under $\sim'$ are unions of fibers under $\sim$.

Theorem 4.6 (Projection Deficit Bound)**.** *Let $\sigma : X \to S$ be a projection inducing fiber partition $\{F_s = \sigma^{-1}(s)\}_{s \in S}$. Let X carry a probability measure μ. Then the deficit increase satisfies:*

$$\Delta\delta(\sigma) \;\geq\; H(X \mid S) \;=\; -\sum_{s \in S} \mu(F_s) \sum_{x \in F_s} \frac{\mu(\{x\})}{\mu(F_s)} \log \frac{\mu(\{x\})}{\mu(F_s)},$$

the conditional entropy of X given S.

Proof. Describing $x \in X$ via $\sigma(x) \in S$ plus a within-fiber index requires at least $H(X \mid S)$ additional bits beyond the description of $\sigma(x)$, by Shannon's source coding theorem. Projection discards this index, making $H(X \mid S)$ bits of distinguishing information inaccessible. Hence $\Delta\delta(\sigma) \geq H(X \mid S)$. □

Corollary 4.7. *When fibers are uniform of size* $|F_s|$*, the bound reduces to* $\Delta\delta(\sigma) \geq \log|F_s|$*, recovering the earlier finite-fiber result as a special case.*

The conservation law governing projection is:

$$|H_t| + |\Omega_t| = |\Omega_0|, \tag{1}$$

where H_t is the realized history, Ω_t the remaining possibility space, and $|\cdot|$ measures information content. Projection transfers possibility from the accessible history into the inaccessible remainder.

4.4. Embedding

Definition 4.8 (Embedding)**.** An *embedding* of $(X, \sim, T)$ into $(Y, \sim', T')$ is an injective map $\phi : X \to Y$ such that the fiber structure is preserved: $\phi(x) \sim' \phi(y) \implies x \sim y$. An embedding is *faithful* if the converse also holds: $x \sim y \iff \phi(x) \sim' \phi(y)$.

Theorem 4.9 (Embedding Refinement)**.** *Let* $\phi : (X, \sim, T) \to (Y, \sim', T')$ *be a faithful embedding. Then:*

$$\delta_{T'}(\phi(x)) \;\leq\; \delta_T(x).$$

A faithful embedding never increases the coding deficit.

Proof. We need to show $L_{T'}(\phi(x)) - L^*(\phi(x)) \leq L_T(x) - L^*(x)$.

Since ϕ is faithful, every description of x reachable in T corresponds to a description of $\phi(x)$ reachable in T': the image ϕ preserves the full fiber structure, so $L_{T'}(\phi(x)) \leq L_T(x)$.

For the optimal lengths: ϕ is an injective map, so any optimal description of $\phi(x)$ can be decoded to an optimal description of x by applying ϕ^{-1} to the fiber structure. This means $L^*(\phi(x)) \geq L^*(x)$: the optimal description of $\phi(x)$ cannot be shorter than the optimal description of x, because the description must encode at least as much distinguishing information (the embedding may add structure in Y that is extrinsic to x, making optimal descriptions of $\phi(x)$ longer, not shorter).

Combining: $\delta_{T'}(\phi(x)) = L_{T'}(\phi(x)) - L^*(\phi(x)) \leq L_T(x) - L^*(x) = \delta_T(x)$. □

Remark 4.10. The key inequality is $L^*(\phi(x)) \geq L^*(x)$, not $\leq$. An injective embedding into a richer space means a description of the image $\phi(x)$ must specify $\phi(x)$'s position within Y, which contains X as a proper subset. This can only make optimal descriptions longer (or equal), not shorter. Non-faithful embeddings that collapse fibers may admit shorter optimal descriptions for the image, but then they are not embeddings in the present sense.

4.5. Revision

Definition 4.11 (Revision)**.** An *ontological revision* is a transition $(X, \sim_T, T) \to (X, \sim_{T'}, T')$ where $T' \supset T$ (every description reachable in T is reachable in T'). A revision is *productive* if $\delta_{T'}(x) < \delta_T(x)$.

Theorem 4.12 (Revision Monotonicity)**.** *Let $T_0 \subset T_1 \subset T_2 \subset \cdots$ be an expanding sequence of ontologies. Then for all $n \geq 0$:*

$$\delta_{T_{n+1}}(x) \;\leq\; \delta_{T_n}(x).$$

Proof. Since $T_n \subset T_{n+1}$, every description reachable in T_n is reachable in T_{n+1}, so $L_{T_{n+1}}(x) \leq L_{T_n}(x)$. The optimal length $L^*(x) = K(x)$ is independent of the ontology. Hence $\delta_{T_{n+1}}(x) = L_{T_{n+1}}(x) - K(x) \leq L_{T_n}(x) - K(x) = \delta_{T_n}(x)$. □

Corollary 4.13 (Asymptotic deficit)**.** *The sequence $(\delta_{T_n}(x))_{n\geq 0}$ is monotone non-increasing and bounded below by zero. Therefore:*

$$\lim_{n\to\infty} \delta_{T_n}(x) \;=\; \delta_\infty(x) \;\geq\; 0.$$

The residual deficit $\delta_\infty(x)$ *exists for every revision sequence. $\delta_\infty(x) = 0$ iff there exists a productive revision sequence converging to an optimal ontology for x.*

The canonical failure of revision is *ontological lock-in*: $\delta_T(x)$ is large and $\delta_\infty(x) > 0$, meaning no revision sequence reachable from within T can achieve optimality for x.

4.6. Transport

Definition 4.14 (Transport)**.** A *transport map* is a function $\tau : (X, \sim_X, T_X) \to (Y, \sim_Y, T_Y)$ with distortion at most $\epsilon \geq 0$:

$$x \sim_X x' \;\implies\; \tau(x) \sim_Y \tau(x'),$$

and the measure of pairs $(\tau(x), \tau(x'))$ with $\tau(x) \sim_Y \tau(x')$ but $x \not\sim_X x'$ is controlled by ϵ.

Theorem 4.15 (Transport Stability)**.** *Let $\tau : (X, \sim_X, T_X) \to (Y, \sim_Y, T_Y)$ be a transport map with distortion at most ϵ. Let $I_X : X \to \mathbb{R}$ be an invariant constant* *on fibers of* $\sim_X$*, and let* $I_Y : Y \to \mathbb{R}$ *be a corresponding invariant constant on fibers of* $\sim_Y$ *with* $I_Y(\tau(x))$ *well-defined for all* $x \in X$*. Suppose* I_X *and* I_Y *agree on perfectly transported fibers:* $I_X(x) = I_Y(\tau(x))$ *whenever* τ *maps the fiber of* x *in* X *isomorphically to the fiber of* $\tau(x)$ *in* Y*. Then:*

$$|I_X(x) - I_Y(\tau(x))| \;\leq\; C\epsilon$$

for a constant C *depending on the oscillation of* I_X *and* I_Y *across fiber boundaries.*

Proof. Let $E_\epsilon \subseteq X$ be the set of x for which τ violates fiber preservation (either collapsing a distinction or creating a spurious one). By the distortion bound, $\mu(E_\epsilon) \leq \epsilon$ for some natural measure μ on X. For $x \notin E_\epsilon$, the fiber of x maps isomorphically to the fiber of $\tau(x)$, so $I_X(x) = I_Y(\tau(x))$ by assumption.

For $x \in E_\epsilon$, the invariants can differ by at most the oscillation of I_X (or I_Y) over a single fiber: $|I_X(x) - I_Y(\tau(x))| \leq \mathrm{osc}(I)$. Taking $C = \mathrm{osc}(I)/\min(\mu(F))$ where the minimum is over non-trivial fibers, the bound $|I_X(x) - I_Y(\tau(x))| \leq C\epsilon$ holds uniformly. □

Corollary 4.16 (Justification of analogy)**.** *An analogical inference transported by* τ *is valid up to error* $C\epsilon(\tau)$*. An analogy is* justified *for inference purposes iff* $\epsilon(\tau)$ *is controlled and the relevant invariants have bounded oscillation. Transport Stability is the precise mathematical condition under which analogical reasoning is valid.*

5. Asymmetries, Failure Modes, and the Meta-Theorem

5.1. Projection and Embedding Are Not Duals

Both are distinction-level operations, but they are not inverses. Projection necessarily has $\Delta\delta(\sigma) \geq H(X \mid S) > 0$ for non-trivial fibers: it always loses information. Embedding has $\Delta\delta(\phi) \leq 0$ for faithful embeddings but may introduce spurious distinctions ($\epsilon(\phi) > 0$) for non-faithful ones. Neither recovers what the other destroys in the general case.

5.2. The Diagnostic Grid

Each failure mode yields a testable diagnosis:

(1) *Projection failure*: $H(X \mid S)$ is large. Distinctions needed for the task have been collapsed. Return to a less-projected representation.

(2) *Embedding failure*: $\epsilon(\phi) > 0$ and spurious distinctions are driving behavior. Coarsen the embedded space.

(3) *Revision failure*: $\delta_\infty(x) > 0$. The residual deficit is positive; the system is stuck. Identify what template expansion would trigger a productive revision.

(4) *Transport failure*: $\epsilon(\tau)$ is large relative to the oscillation bound in Theorem 4.15. The analogy is invalid; quantify and correct the distortion.

5.3. The Meta-Theorem

Theorem 5.1 (Meta-theorem on representational adequacy)**.** *A representational system manages its ontological deficit well iff:*

(i) Projection*: $H(X \mid S)$ is small for task-relevant objects.*

(ii) Embedding*: $\epsilon(\phi)$ is controlled; few spurious distinctions.*

(iii) Revision*: $\delta_\infty(x) = 0$; residual deficit vanishes under the reachable revision sequence.*

(iv) Transport*: $C\epsilon(\tau)$ is within acceptable tolerance for invariants of interest.*

Each condition is independently measurable via conditional entropy, false discovery rate, asymptotic deficit convergence, and transport distortion respectively.

6. The Deficit as Universal Invariant

6.1. Portability

Compression. The deficit is the redundancy of archive T for object x. The Revision Monotonicity theorem guarantees that a productive revision sequence converges to a residual deficit, and the compression jumps are exactly the productive revisions.

Machine learning. The deficit is approximation error: the component of generalization error due to hypothesis class limitation that more data cannot reduce.

Scientific discovery. Normal science is the regime of small, decreasing δ_T within the current paradigm. Scientific revolution is a productive revision with large $\Delta\delta(\rho) < 0$: the paradigm shift reaches a new T' with much lower deficit for the anomalous observations.

Language evolution. New terms are productive revisions of the linguistic archive; each coinage decreases $\delta_T(x)$ for objects previously requiring circumlocution.

Repair theory. Repair is transport from a high-deficit damaged state to a low-deficit repaired state; the validity of the repair is governed by Transport Stability.

6.2. Subadditivity

Theorem 6.1 (Deficit Subadditivity)**.** *Let $x \oplus y$ denote an object whose optimal description decomposes independently into descriptions of x and y. Then:*

$$\delta_T(x \oplus y) \;\le\; \delta_T(x) + \delta_T(y) + O(\log n),$$

where $n = |x| + |y|$ is the combined description length and the $O(\log n)$ term accounts for the overhead of specifying the decomposition.

Proof. $L_T(x \oplus y) \le L_T(x) + L_T(y)$ since a description of $x \oplus y$ can be formed by concatenating descriptions of x and y reachable in T. By near-additivity of Kolmogorov complexity for independent objects, $K(x \oplus y) \ge K(x) + K(y) - O(\log n)$. Hence $\delta_T(x \oplus y) = L_T(x \oplus y) - K(x \oplus y) \le (L_T(x) + L_T(y)) - (K(x) + K(y) - O(\log n)) = \delta_T(x) + \delta_T(y) + O(\log n)$. □

Remark 6.2. This mirrors the subadditivity of Kolmogorov complexity and makes the deficit behave like a legitimate information quantity. The Deficit Subadditivity Theorem connects the framework directly to coding theory: the deficit of a composed system is no worse than the sum of deficits of its parts.

7. Toward a Categorical Formulation

The categorical formulation must be derived, not imported. We proceed in strict order: objects, morphisms, operations, functors.

7.1. The Category Dist

Definition 7.1 (The category **Dist**)**.** The category **Dist** has:

- *Objects*: ontological triples $(X, \sim, T)$.
- *Morphisms*: functions $f : (X, \sim_X, T_X) \to (Y, \sim_Y, T_Y)$ with associated distortion $\epsilon(f) \ge 0$ and deficit change $\Delta\delta(f) \in \mathbb{R}$.

- *Identity*: the identity map on $(X, \sim, T)$ with $\epsilon = 0$, $\Delta\delta = 0$.
- *Composition*: $(g \circ f)$ has $\epsilon(g \circ f) \leq \epsilon(f) + \epsilon(g)$ (distortions accumulate) and $\Delta\delta(g \circ f) = \Delta\delta(f) + \Delta\delta(g)$ (deficit changes add).

By including T in the object, the deficit δ_T is a genuine invariant of objects, not an additional datum to be tracked separately. This resolves the under-specification present in earlier formulations.

7.2. The Quotient Functor and Projection

Proposition 7.2 (Projection Functor)**.** *Define the quotient functor* $Q : \mathbf{Dist} \to \mathbf{Set}$ *by* $Q(X, \sim, T) = X/\sim$. *Every projection* σ *with* $\sim_1 \subseteq \sim_2$ *induces a quotient map:*

$$Q(\sigma) : X/\sim_1 \to X/\sim_2, \quad [x]_{\sim_1} \mapsto [x]_{\sim_2}.$$

$Q(\sigma)$ *is well-defined and surjective. This identifies projection with quotient formation in* **Dist**.

Proof. Well-definedness: since $\sim_1 \subseteq \sim_2$, every $\sim_1$-class is a subset of a $\sim_2$-class, so the map is consistent. Surjectivity: every $\sim_2$-class is the image of some $\sim_1$-class. □

7.3. Morphism Types and the Hierarchy

The two-level hierarchy maps onto morphism types:

Projection: endomorphism with $\Delta\delta \geq 0$, Q-image a quotient map.

Embedding: morphism into a larger object $(Y, \sim', T')$ with $Y \supset X$, $\Delta\delta \leq 0$ (faithful), $\epsilon \geq 0$ (potential spurious distinctions).

Revision: endomorphism changing T while preserving X, $\Delta\delta < 0$ (productive).

Transport: ordinary morphism between distinct objects, governed by Theorem 4.15.

7.4. The Deficit as Functor

Proposition 7.3 (Deficit functor)**.** *Define* $\delta : \mathbf{Dist} \to \mathbf{Fun}(X, \mathbb{R}_{\geq 0})$ *by* $\delta(X, \sim, T) = (x \mapsto \delta_T(x))$. *On morphisms,* $\delta(f)$ *is the function* $\delta_T(x) \mapsto \delta_{T'}(f(x)) = \delta_T(x) + \Delta\delta(f)$. *This assignment is functorial:* $\delta(\mathrm{id}) = \mathrm{id}$ *and* $\delta(g \circ f) = \delta(g) \circ \delta(f)$.

Remark 7.4. This is a well-defined functor because δ_T depends on T, which is part of the object. The action on morphisms is determined by the signed deficit change $\Delta\delta(f)$, which is part of the morphism data. The natural transformation interpretation from earlier drafts holds: δ measures morphism quality, and perfect morphisms are those with $\Delta\delta = 0$, $\epsilon = 0$.

7.5. Relation to Admissibility Geometry

The admissibility relation $\mathcal{A}$ specifies the subcategory of **Dist** whose morphisms have $\epsilon(f) \leq \epsilon_{\max}$ and $\Delta\delta(f) \leq \Delta_{\max}$. In the RSVP framework, inadmissible transitions are those that increase deficit or distortion beyond acceptable thresholds. The categorical structure of **Dist** subsumes admissibility geometry as the subcategory determined by these bounds, connecting the formal framework to the broader research programme.

8. Conclusion

Distinguishability geometry proposes the ontological triple $(X, \sim, T)$ as the primitive object for the study of representational change. The deficit is the invariant. The four operations are the dynamics. The Classification Conjecture is the central open problem.

What is established. The Deficit Proxy Theorem makes the coding deficit a computable surrogate for the geometric distinguishability deficit. The Representational Reduction Proposition establishes the four operations as the exhaustive qualitative fiber effects. Each operation has a characteristic theorem. The Distortion–Deficit Relationship connects the two error measures. The Meta-Theorem restates representational adequacy as four independently measurable conditions. The category **Dist** is well-defined and the deficit is a genuine functor.

What remains open.

(1) *Full Deficit Correspondence* (Conjecture 2.9): does $\Delta_T(x) = 0 \iff \delta_T(x) = 0$?

(2) *Classification Conjecture* (Conjecture 3.1): does every representational transformation factor into a finite composition of the four operations?

(3) *Residual deficit*: when does $\delta_\infty(x) = 0$? What structural property of the revision sequence ensures convergence to zero?

(4) *Transport metric*: is there a canonical metric on **Dist** making transport maps Lipschitz?

(5) *Two-level duality*: is there a duality functor **Dist** $\to$ **Dist**$^{\mathrm{op}}$ exchanging projection with transport and embedding with revision?

(6) *Unified deficit bound*: does a single expression govern $\delta_{T'}(x) \leq \delta_T(x) + C(\sigma, \phi, \rho, \tau)$ for arbitrary composed transformations?

The strongest reading of the paper is not "here is a complete theory" but "here is a candidate primitive ontology and four conjectured primitive operations." These six open problems define exactly where the theory consolidates or breaks. They are where the work is.

A. Formal Definitions: Fiber Bundles and the Quotient Construction

An ontological triple $(X, \sim, T)$ corresponds to a fiber bundle $E \to B$ where $B = X/{\sim}$ is the base space of equivalence classes, each fiber $q^{-1}(b) = [x]_\sim$ lies above its class, and the ontology T equips E with a description system. Projection increases fiber size (bundle coarsening). Embedding decreases it (bundle refinement). Revision changes B by changing T. Transport maps between distinct bundles.

The deficit $\delta_T(x)$ is an invariant of the triple, not of the bundle alone: two triples $(X, \sim, T)$ and $(X, \sim, T')$ with the same fiber structure but different ontologies carry different deficit functions.

B. The Conservation Law

In the event-sourced setting: at $t = 0$, $|\Omega_0| = \log|\mathrm{supp}(\mu)|$. As history h_t is realized, $|H_t| + |\Omega_t| = |\Omega_0|$ because information transferred to actuality is subtracted from possibility. The Projection Deficit Bound (Theorem 4.6) is consistent: the $H(X \mid S)$ bits of information destroyed by projection move from accessible history into the inaccessible remainder Ω_t.

C. Deficit Examples and the Finite Illustration

Compression. LZ over periodic string: $L_T(x) \approx |x|/\log|x|$, $K(x) \approx \log p + \log(|x|/p)$ for period p. Deficit is bits wasted before the archive learns the period.

clio projection error. For history h with $\sigma(h) = s$ and fiber F_s, $\delta_T(h) = H(X \mid S) \geq \log |F_s|$ consistent with Theorem 4.6.

Machine learning. Linear class over nonlinear data: the deficit is the approximation error floor. Revision Monotonicity guarantees that expanding the hypothesis class (a productive revision) decreases this floor.

Finite illustration revisited. In Example 2.11 with $X = \{a, b, c, d\}$: the projection σ has $|F_p| = 2$, so $\Delta\delta(\sigma) \geq \log 2 = 1$ bit. The revision from T_0 to T_1 has $\delta_{T_1}(a) = \delta_{T_1}(b) = 0 < \delta_{T_0}(a) = \delta_{T_0}(b)$ when working from the projected space: a productive revision. The transport $\tau : X \to Y$ has $\epsilon(\tau) = 0$ and hence $\Delta\delta(\tau) = 0$ by Theorem 4.3: a perfect transport.