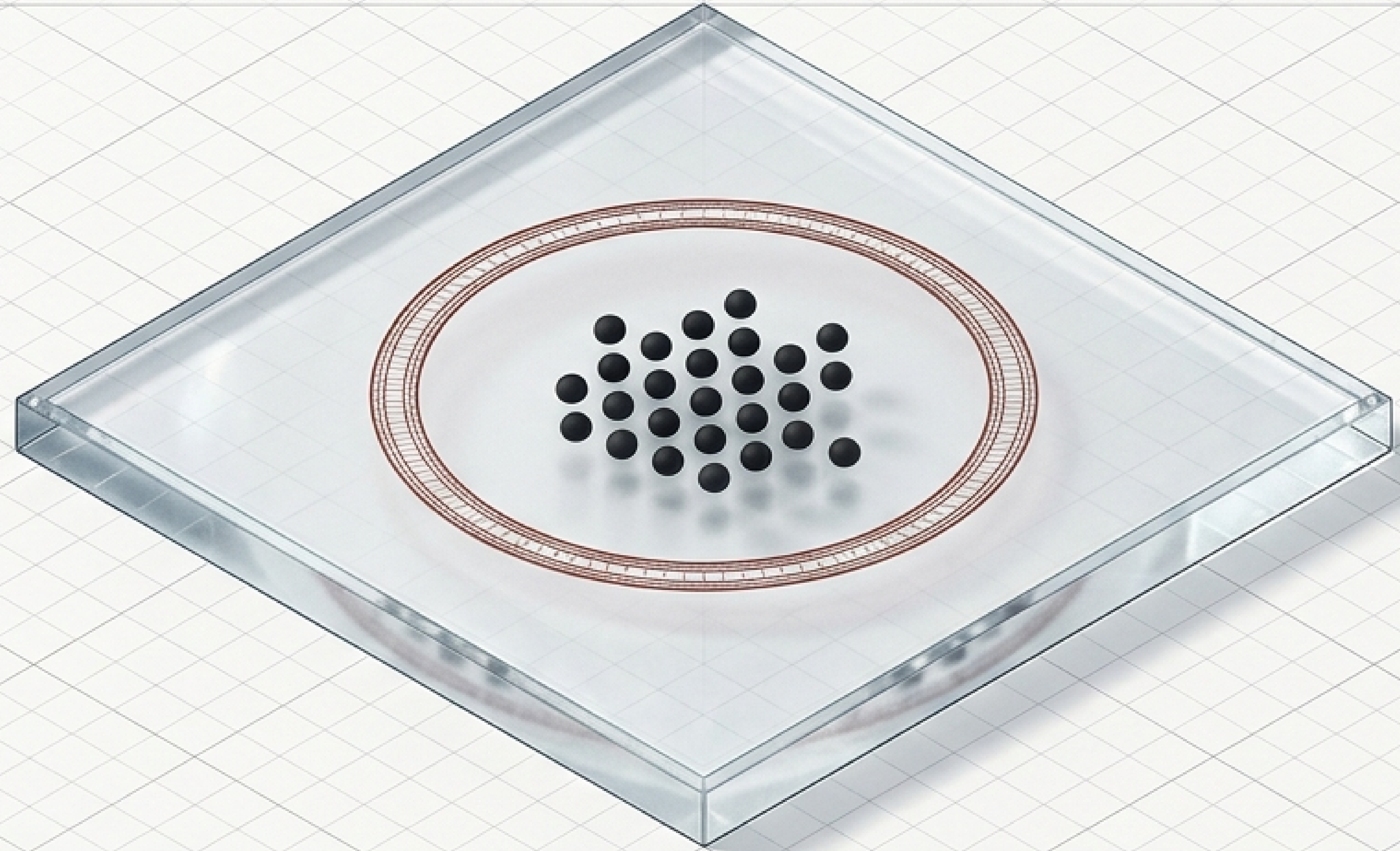
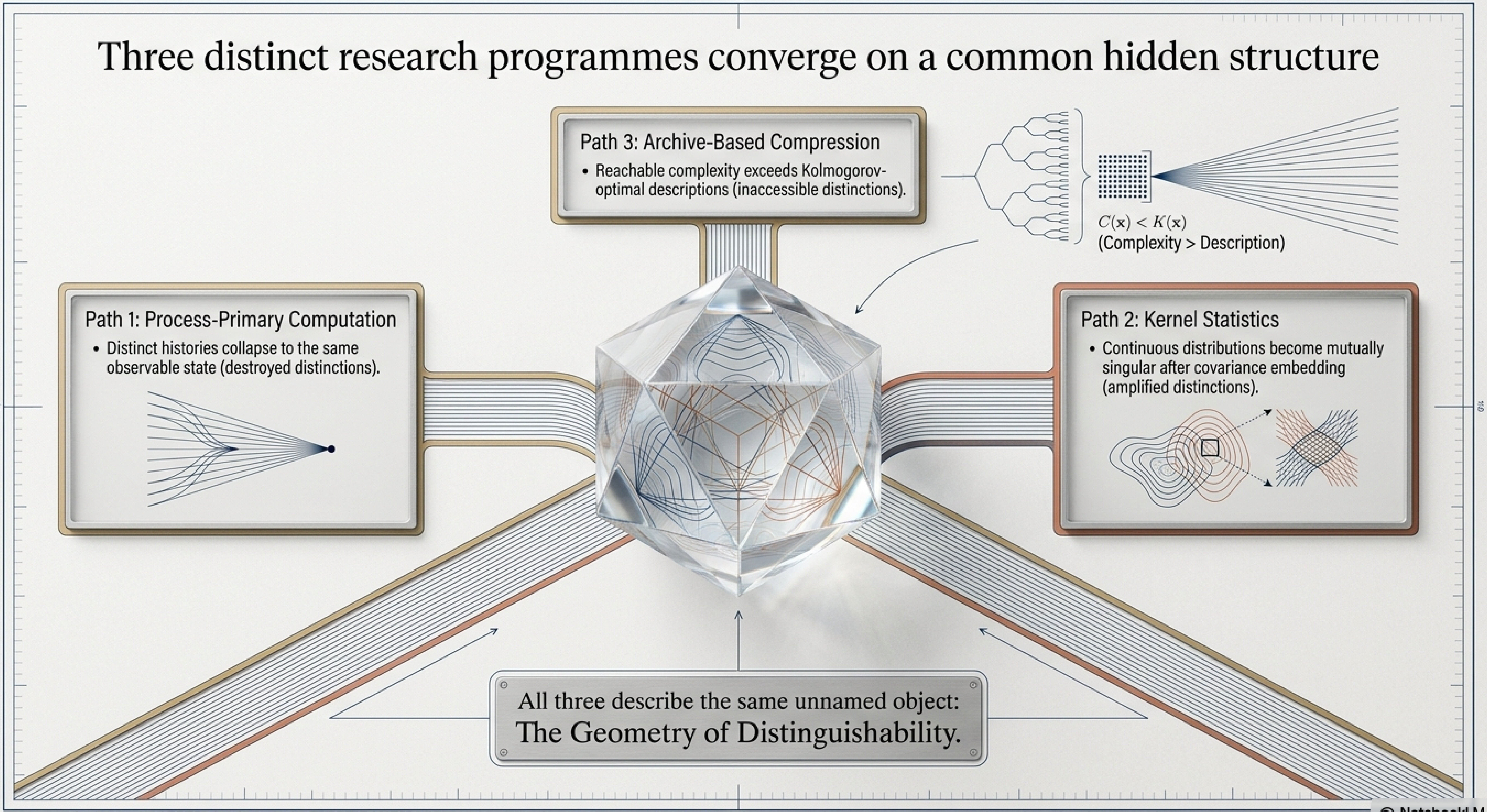
The Geometry of Distinguishability

How representational systems gain, lose, and transfer the capacity to express distinctions.



Based on the mathematical framework of projection, embedding, revision, and transport.


Three distinct research programmes converge on a common hidden structure
Path 3: Archive-Based Compression
• Reachable complexity exceeds Kolmogorov-optimal descriptions (inaccessible distinctions).
$C(\mathbf{x}) < K(\mathbf{x})$
(Complexity > Description)
Path 1: Process-Primary Computation
• Distinct histories collapse to the same observable state (destroyed distinctions).
Path 2: Kernel Statistics
• Continuous distributions become mutually singular after covariance embedding (amplified distinctions).
All three describe the same unnamed object:
The Geometry of Distinguishability.

The foundational primitive is the Distinguishability Space $(X, \sim)$

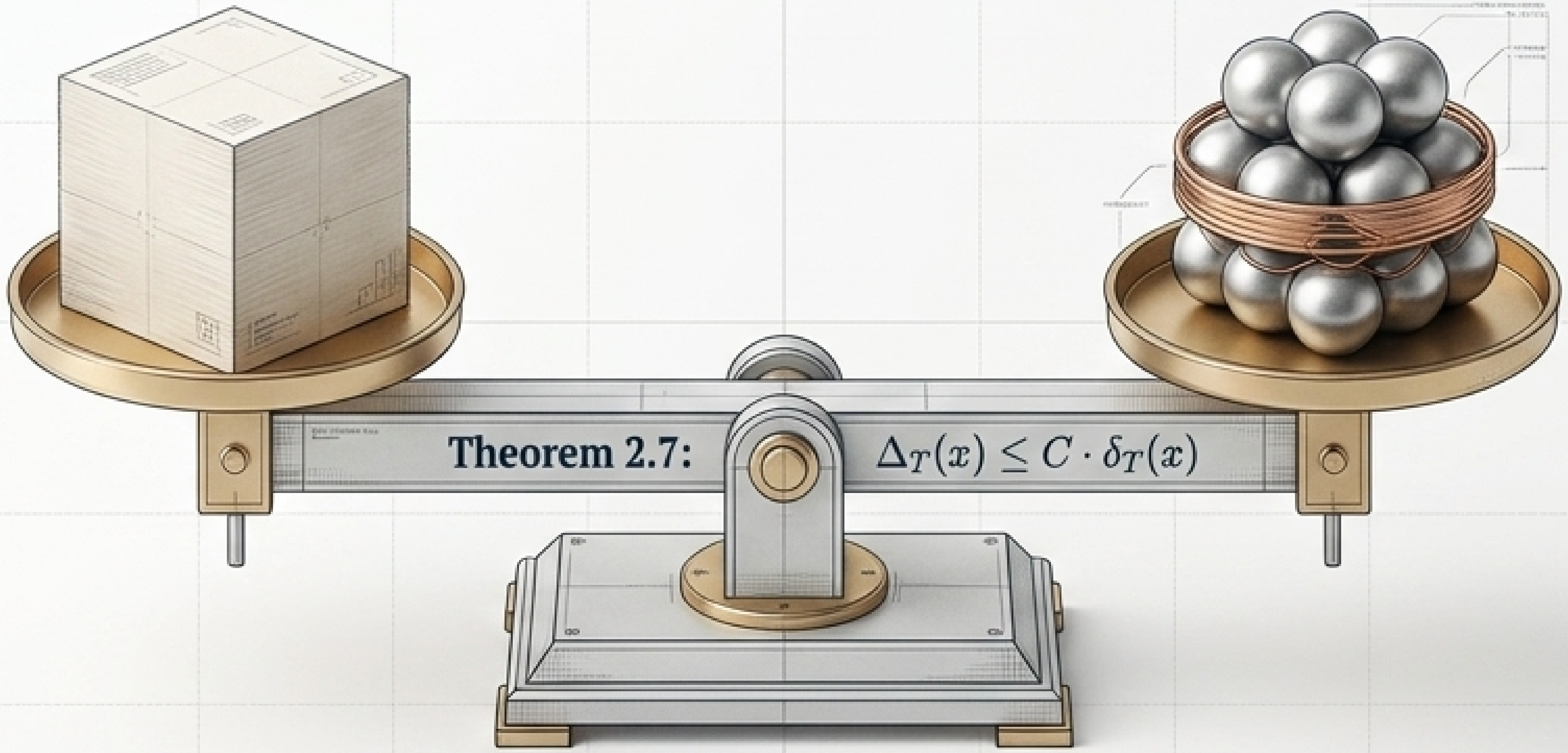
The Distinguishability Principle:
A representational system is characterized not by what objects it contains, but by what distinctions it can express.

- Space (X): The underlying set of elements.
- Relation (~): Observational indistinguishability.
- The Fiber ($[x]_\sim$): The boundary grouping elements that cannot be told apart from x. Finer relations have smaller fibers; coarser relations have larger ones.

The Universal Invariants: Every representational system possesses a deficit

Coding Deficit (δ_T)

- **Concept:** Wasted Bits (Compression).
- **Definition:** $L_T(x) - L^*(x)$ (Description length reachable in Ontology T minus the Kolmogorov-optimal description).

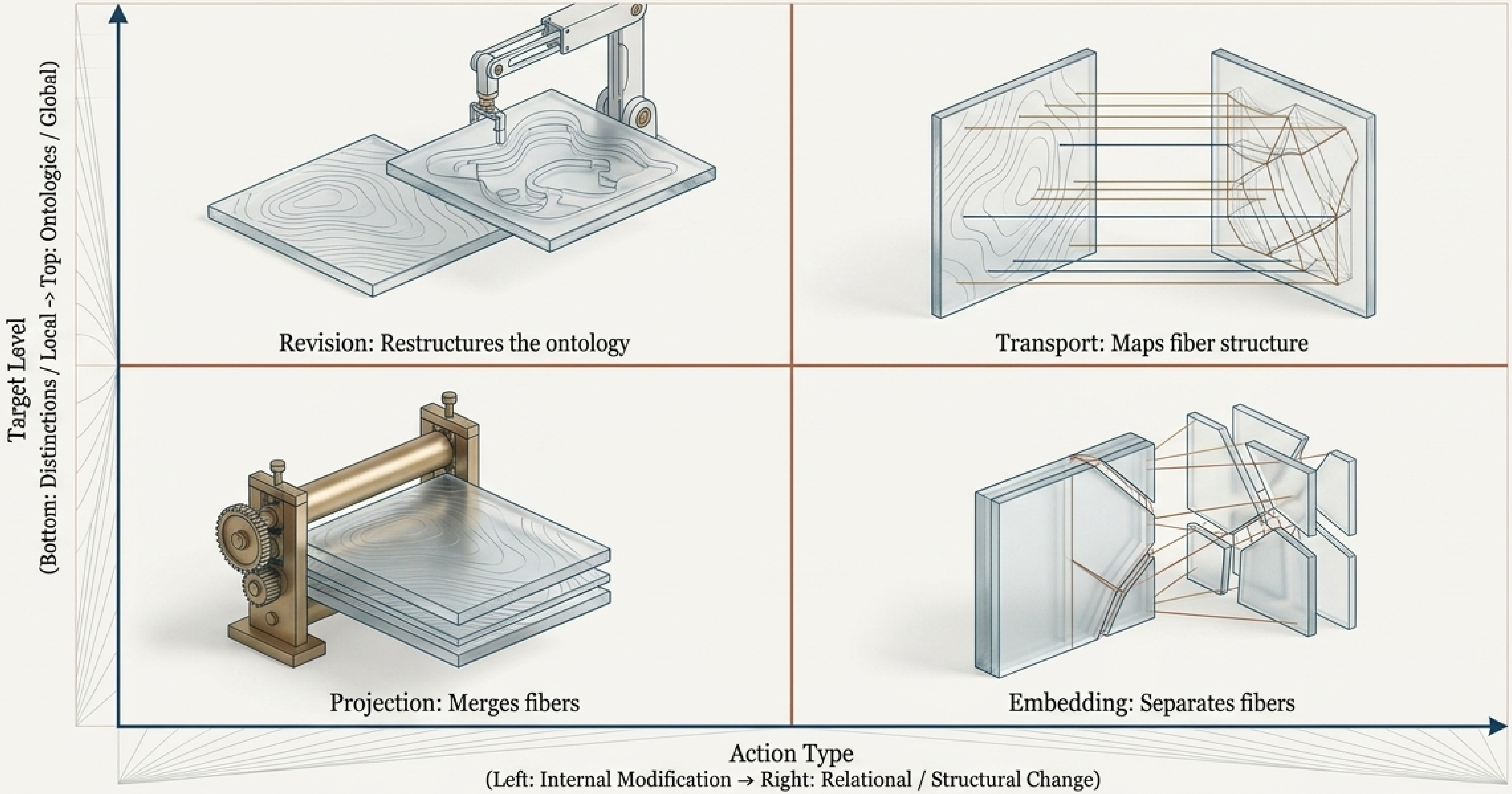


Distinguishability Deficit (Δ_T)

- **Concept:** Hidden Realities (Capacity).
- **Definition:** $D^*(x) - D_T(x)$ (Maximum capacity minus distinctions accessible in T).

Plain-English Meaning: Local optimality in coding corresponds exactly to exhausting all expressible distinctions.

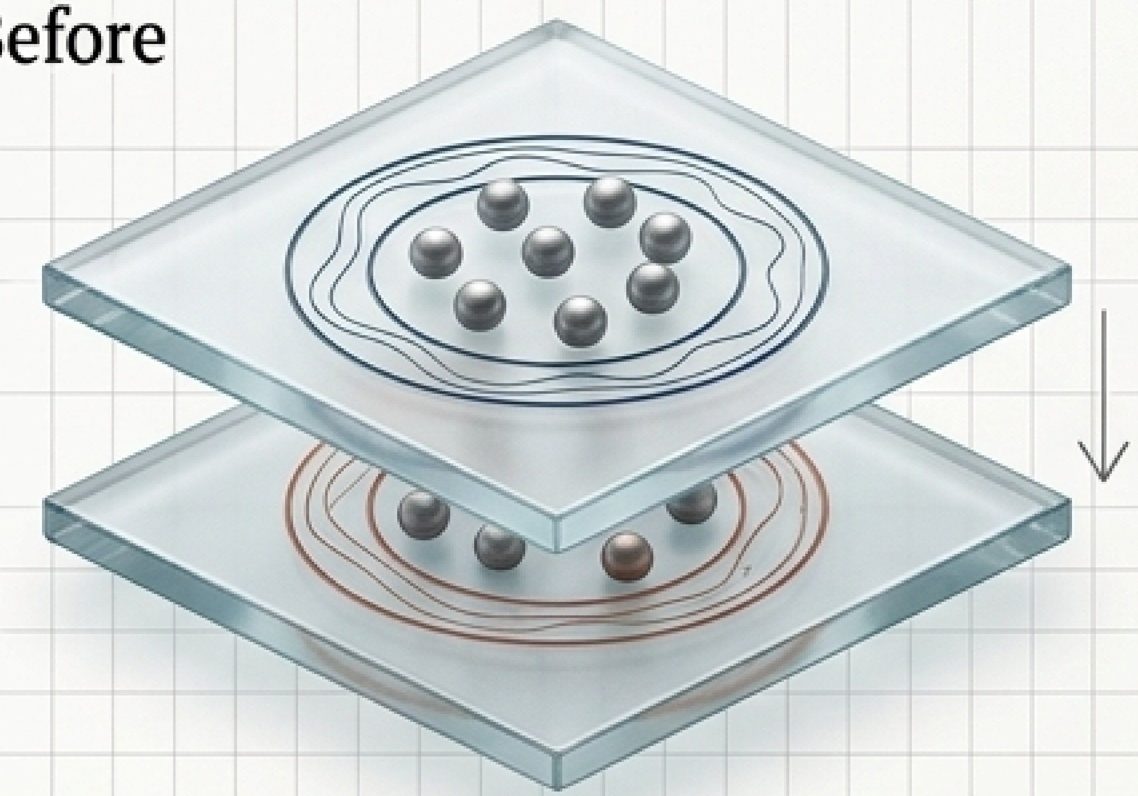
The Two-Level Architecture of Representational Change



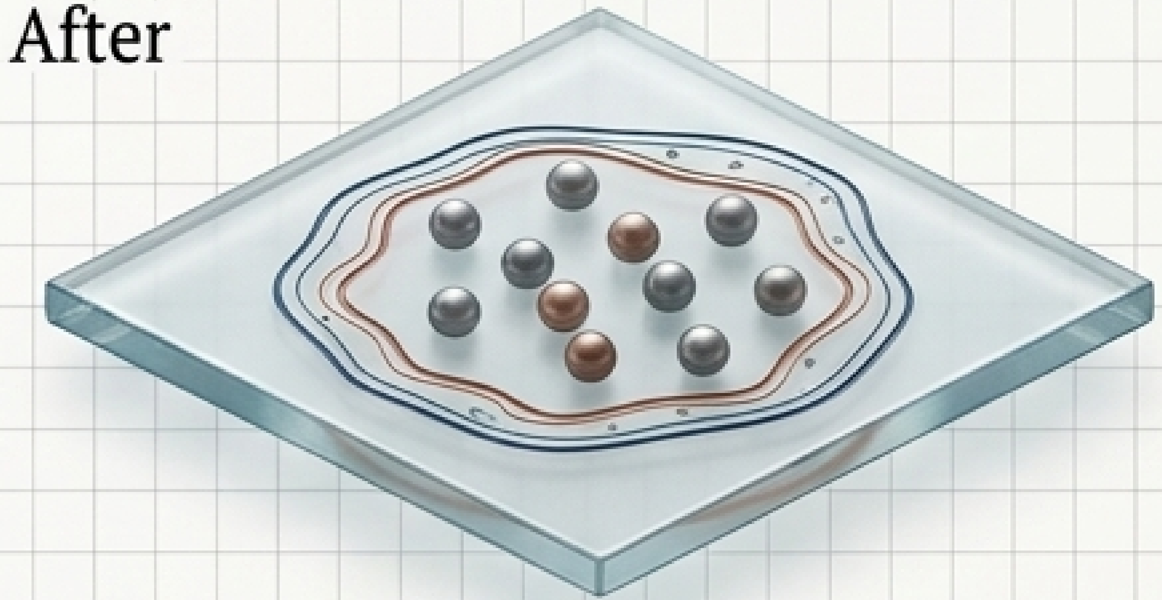
Key Takeaway: These are not an arbitrary taxonomy; they are the exhaustively derived consequences of how fibers can change.

Standardized Operation Layout

Before



After



Operation 1: Projection coarsens relations and destroys distinctions

Effect on $(X, \sim)$:	Merges fibers ($\sim'$ is coarser than $\sim$).
Deficit Shift $(\Delta\delta)$:	≥ 0 (Increases).

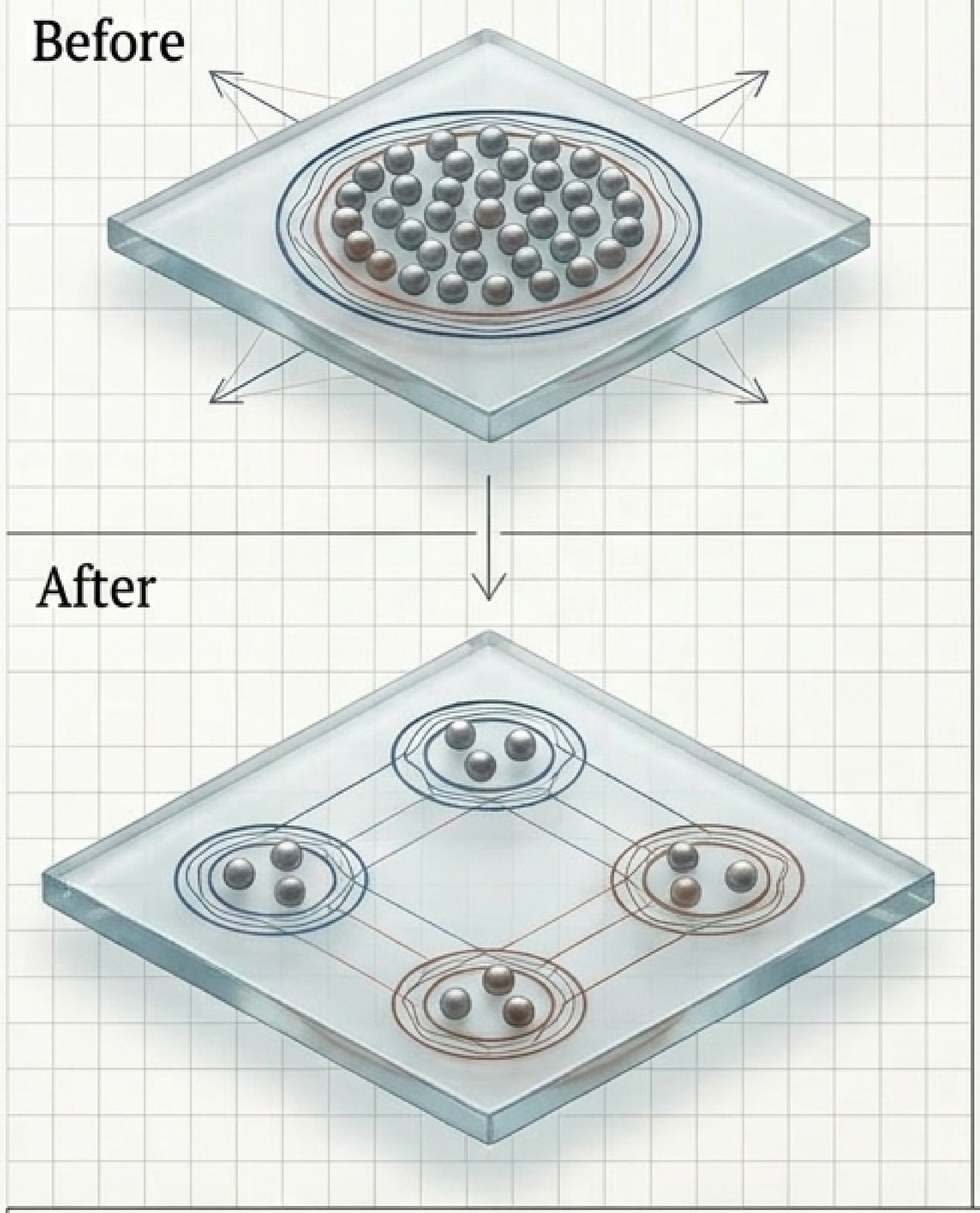
Theorem 3.4 (Projection Deficit Bound):

$$\Delta\delta(\sigma) \geq \log |F_s|$$

Plain-English Meaning:
Selecting an element within a merged fiber requires at least $\log |F_s|$ bits. Projecting discards this index, permanently losing that information.

Conservation Law: $|H_t| + |\Omega_t| = |\Omega_0|$. Possibility is conserved; lost history transfers into the inaccessible remainder.

Before and After



Operation 2: Embedding refines relations and amplifies distinguishability

Effect on $(X, \sim)$:	Separates fibers ($\sim'$ is finer than $\sim$).
Deficit Shift ($\Delta\delta$):	≤ 0 (Decreases, conditionally).

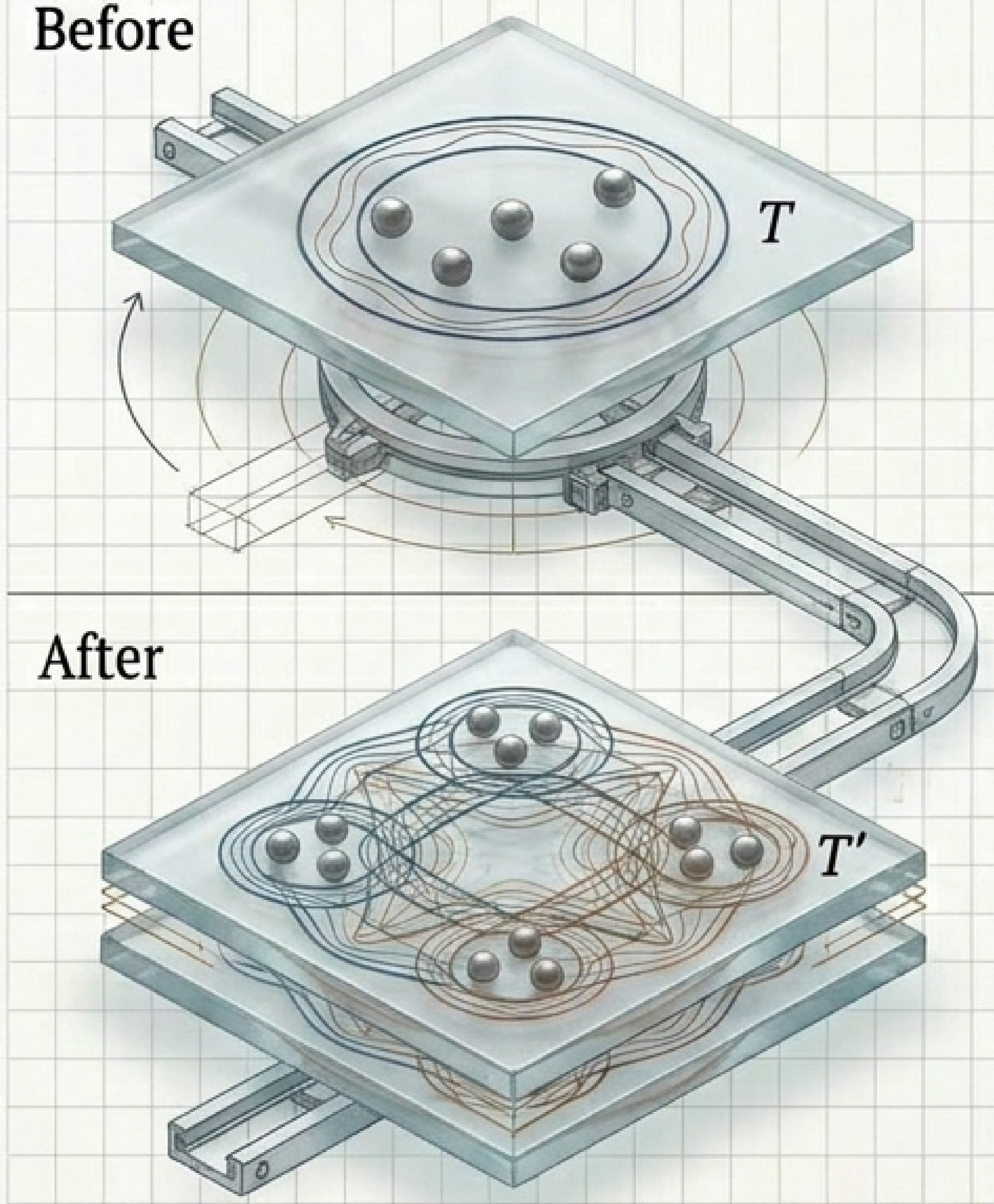
Theorem 3.6 (Embedding Refinement):

$$\delta_{\phi}(X)(\varphi(x)) \leq \delta_X(x)$$

Plain-English Meaning:
A faithful embedding never increases the coding deficit. Every description reachable in X remains reachable in the richer space Y.

The Risk: Non-faithful embeddings introduce spurious distinctions (false positives) measured by distortion $\varepsilon(\phi) > 0$.

Standardized Operation Layout



Operation 3: Revision restructures the ontology to escape lock-in

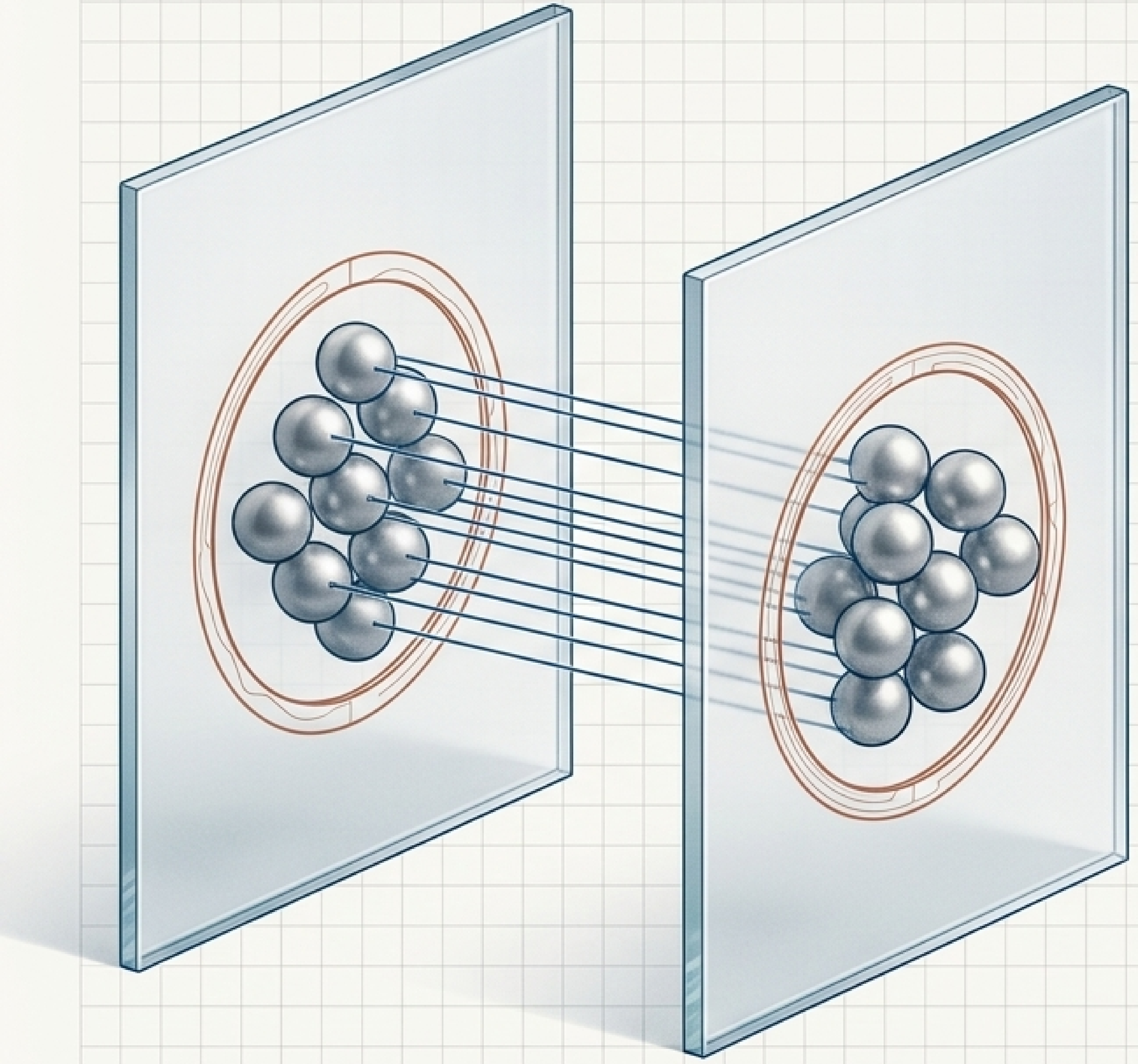
Effect on $(X, \sim)$:	Changes the generating ontology entirely ($T \to T'$).
Deficit Shift ($\Delta\delta$):	< 0 (Decreases under productive revisions).

Theorem 3.9 (Revision Monotonicity):

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

Plain-English Meaning:
A productive revision sequence consistently drives the deficit down. If the deficit hits zero, the system achieves local optimality.

The Failure Mode: Ontological Lock-in. The deficit is positive but stable, and the admissibility path to a better ontology is blocked.



Operation 4: Transport connects distinct spaces, formalizing analogy

Effect on $(X, \sim)$:	Maps $\sim$ across spaces with bounded distortion ε.
Deficit Shift $(\Delta\delta)$:	Bounded by ε.

Theorem 3.12 (Transport Stability):

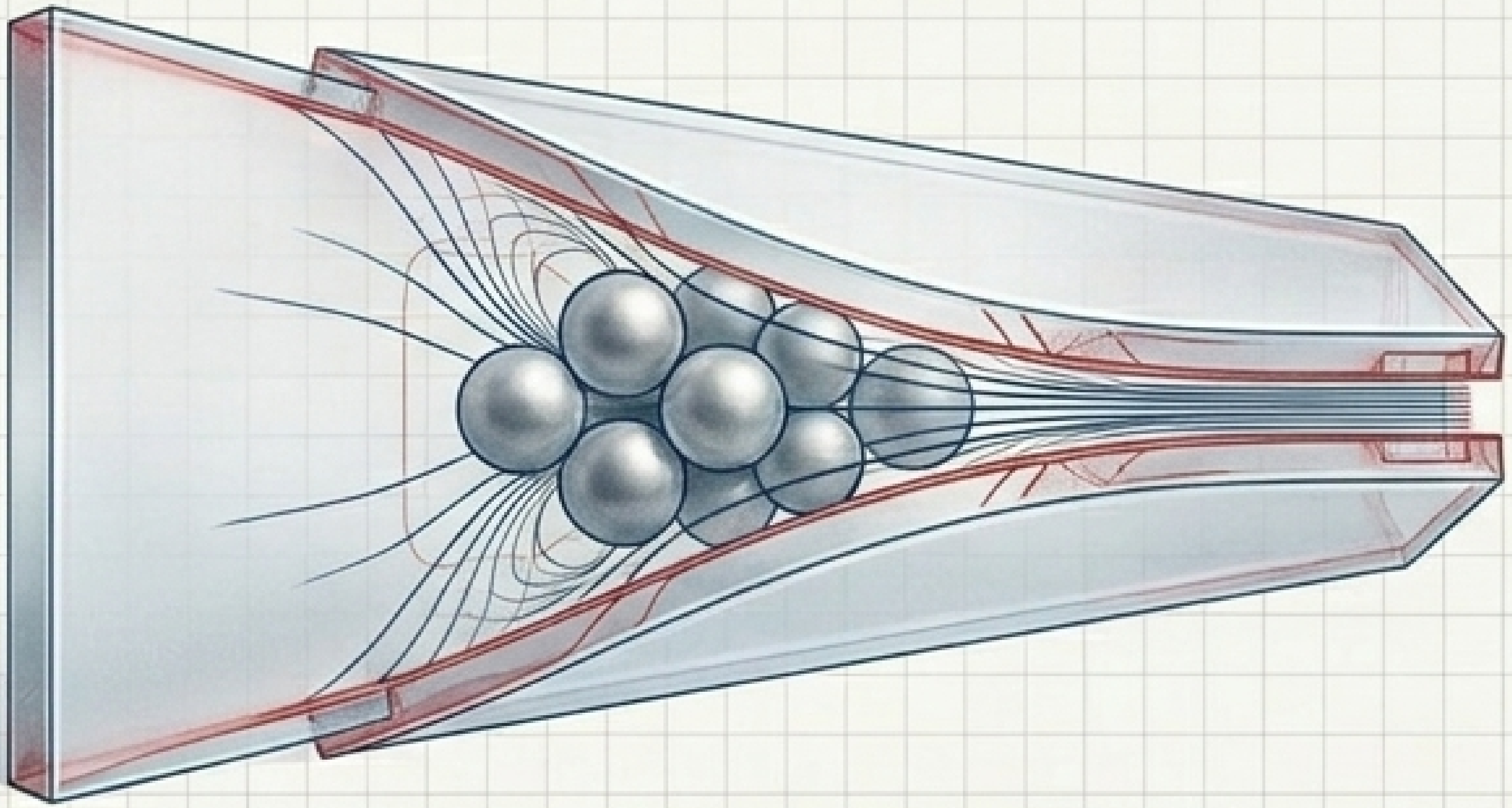
$$|I(\mathbf{x}) - I(\tau(\mathbf{x}))| \leq C\varepsilon$$

Plain-English Meaning: An analogy is mathematically valid for inference if and only if the distortion of the transport map is strictly controlled.

The Insight: Without this stability, analogy is just philosophy. With it, analogy is a quantifiable operation.

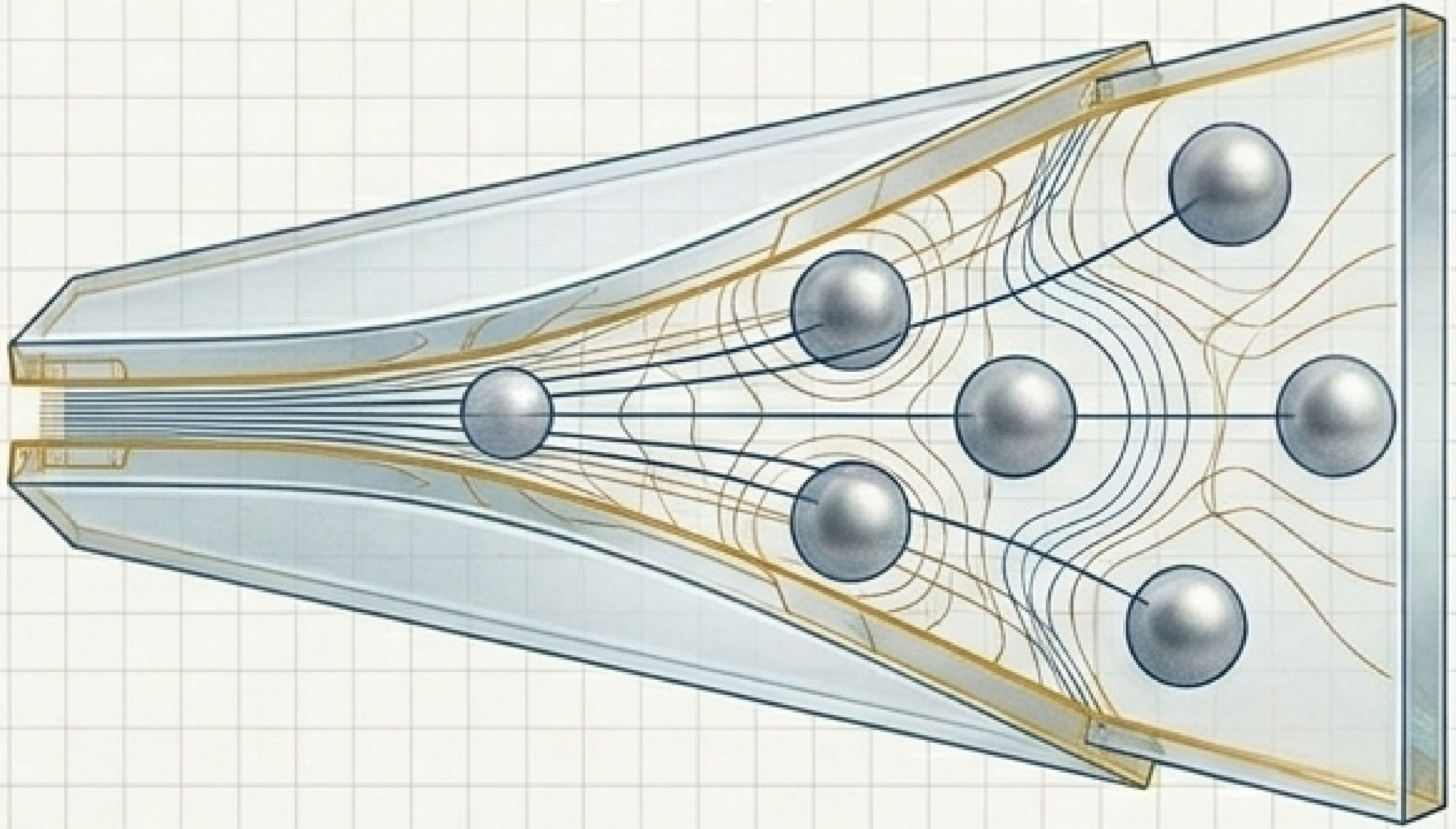
Projection and Embedding are local operations, but they are not duals

The Projection Asymmetry



- **Effect:** Necessarily deficit-increasing.
- **Reality:** Information destroyed by merging fibers is generically unrecoverable from the projected image alone. Real distinctions are lost.

The Embedding Asymmetry



- **Effect:** Conditionally deficit-decreasing.
- **Reality:** While it can reveal hidden structure, it risks introducing spurious distinctions—elements separated in the new space that were equivalent in the old space.

The Diagnostic Grid: Identifying Representational Failure Modes

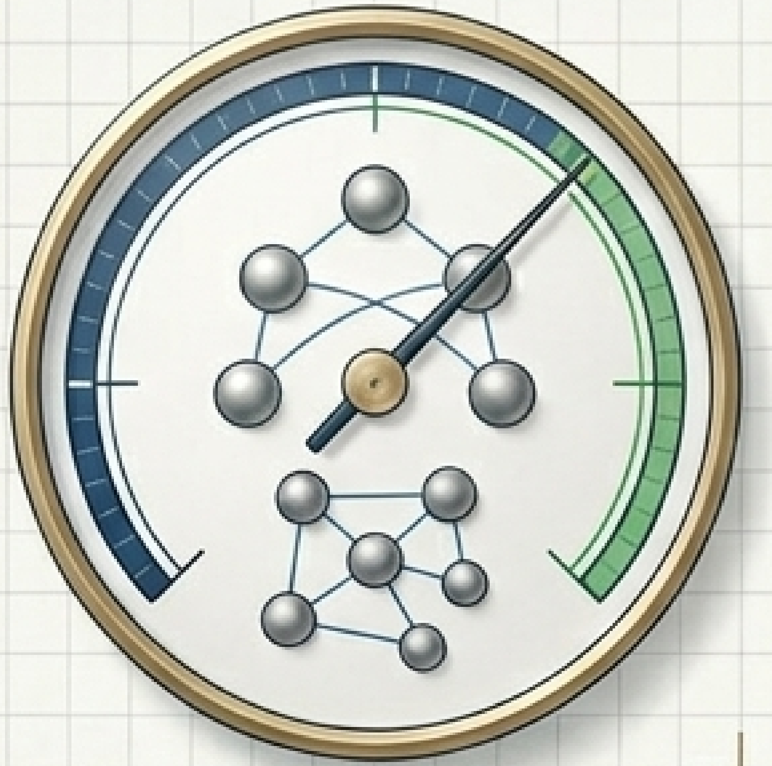
Operation	Level	Deficit (Δδ)	Diagnosis	Fix
Projection	Distinction	≥ 0	Information Loss	Return to a less projected state.
Embedding	Distinction	≤ 0	Distinction Inflation	Coarsen the over-resolved space.
Revision	Ontology	< 0	Ontological Lock-In	Expand the template hierarchy to trigger a jump.
Transport	Ontology	• Bounded	Fiber Distortion	Quantify and correct invalid analogical mappings.

The Meta-Theorem for Representational Adequacy

A system manages its ontological deficit well if and only if these four measurable conditions are met:

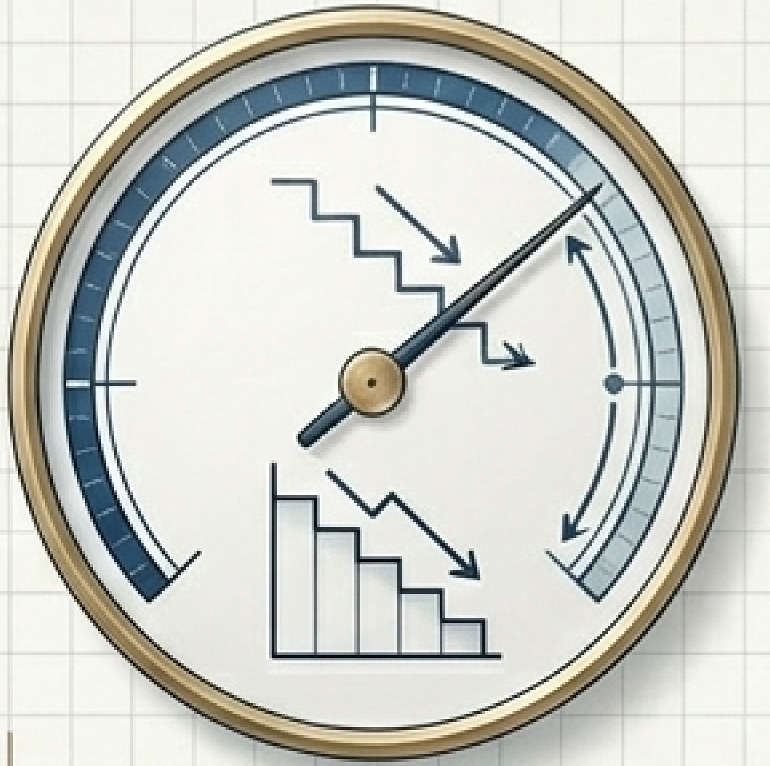
Gauge 1: Preservation under Projection.

Measures information content of projected fibers. Collapsed fibers must not contain task-relevant distinctions.



Gauge 2: Recoverability under Embedding.

Measures false discovery rate. Refinements must correspond to real structure with bounded spurious distinctions.



Gauge 3: Deficit Reduction under Revision.

Measures productivity rate. Revisions must occur before compression plateaus entrench.

Gauge 4: Fidelity under Transport.

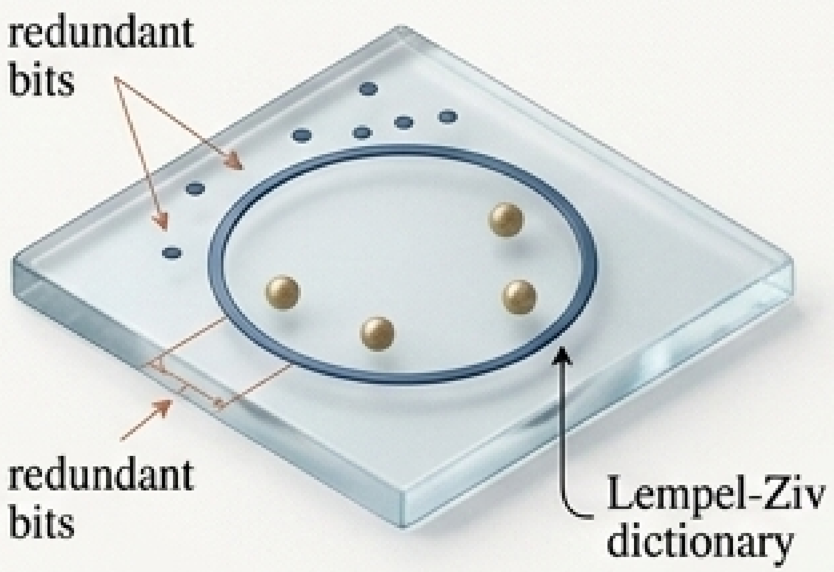
Measures Lipschitz distortion Analogies must preserve task-relevant invariants.

The Classification Conjecture: The Alphabet of Representation

Conjecture 4.2:
Every representational transformation factors into a finite composition of projection, embedding, revision, and transport maps.

The Implication If true, these four operations graduate from a useful taxonomy to the complete set of primitive moves in the geometry of representational change. Every cognitive operation, scientific advance, and computational transformation becomes a word written in the alphabet of $\{\sigma, \phi, \rho, \tau\}$.

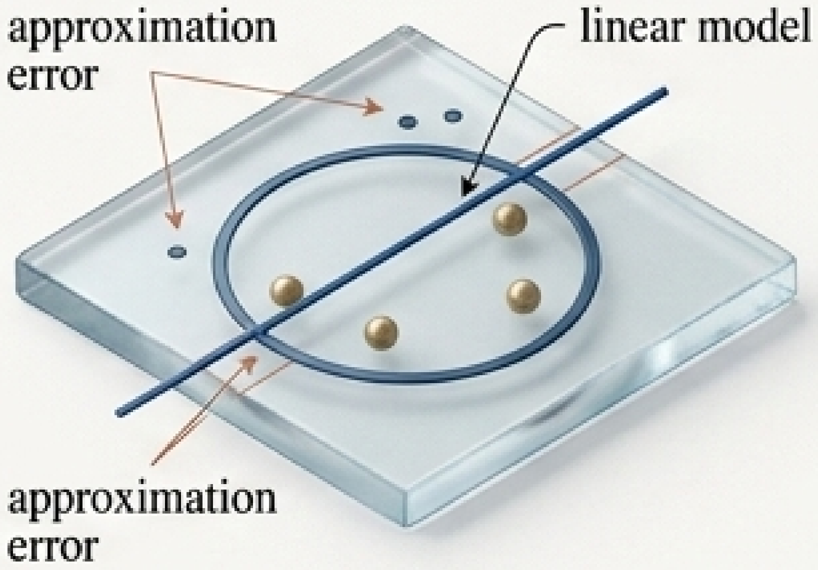
The Deficit is a Universal Invariant Across Domains



Compression:

The Lempel-Ziv dictionary.

Deficit = redundant bits wasted before the archive learns a string's hidden period.

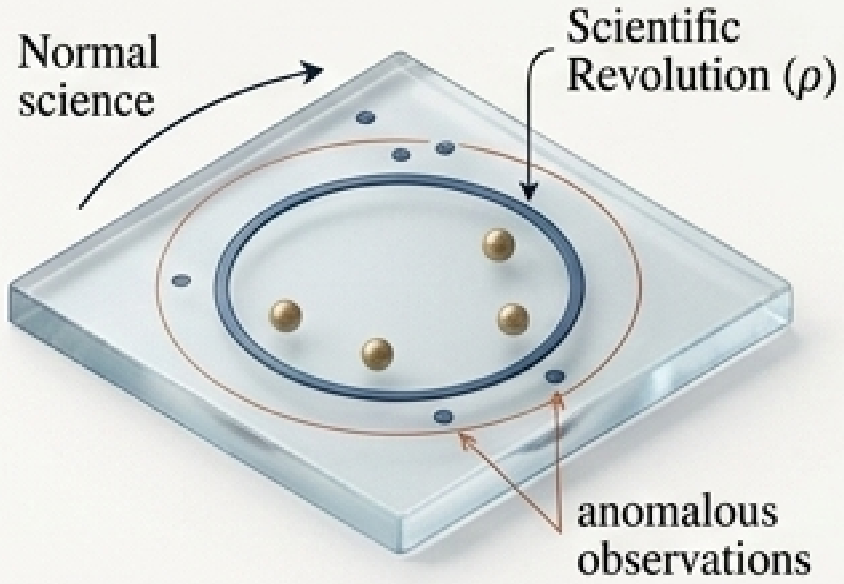


Machine Learning:

A linear model on nonlinear data

Deficit = approximation error

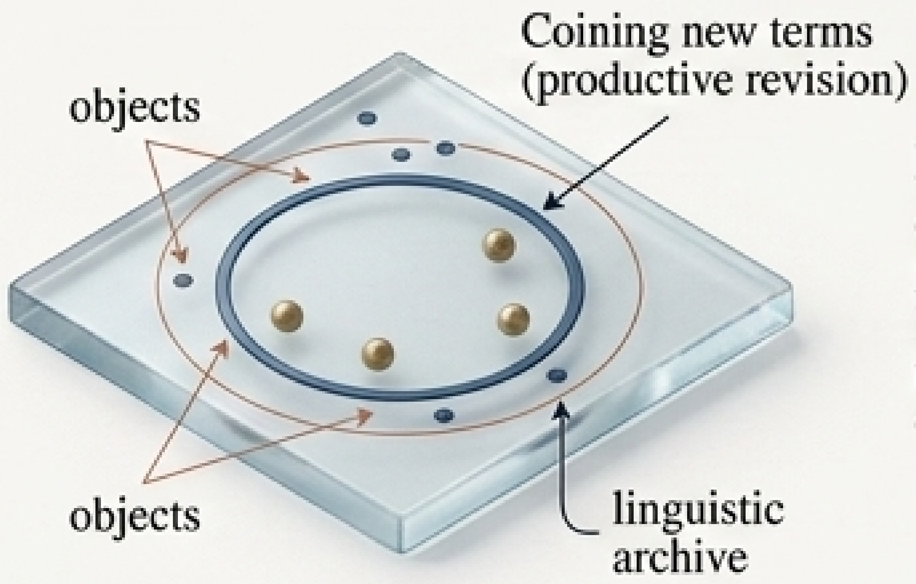
More data cannot reduce it; only a revision of the model class can.



Scientific Discovery:

Normal science = decreasing δ_T

Scientific Revolution = A productive revision (ρ) dropping the deficit for anomalous observations.

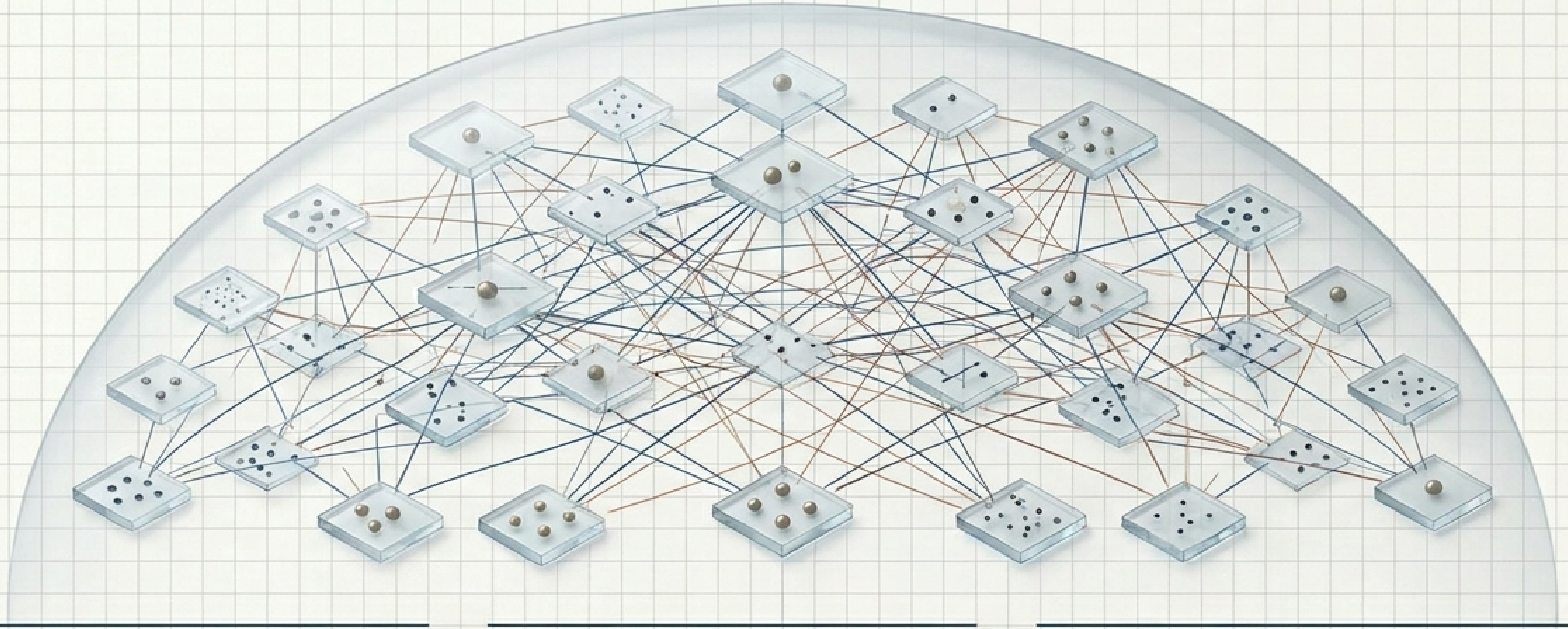


Language Evolution:

Coining new terms = A productive revision of the linguistic archive for objects that previously required circumlocution.

The Categorical Horizon: The Category Dist

The category is not imposed; it naturally emerges from the Distinguishability Principle.



Objects: Distinguishability spaces (X, ~) equipped with their deficit functions.

Morphisms: The four operations form the hierarchy of endomorphisms and morphisms.

The Ultimate Invariant: The Deficit acts as a natural transformation, perfectly measuring the quality, loss, and productivity of every morphism in the category.