

**Persistent Anomalies and the  
Geometry  
of Ontology Revision**

Repair, Admissibility, and Non-Abelian Tears

Flyxion

2026

*An independent research monograph*

# Contents

<b>Preface</b>	<b>vii</b>
<b>I Repair Theory</b>	<b>1</b>
<b>1 Models, Observations, and Residuals</b>	<b>2</b>
1.1 Basic Setup . . . . .	2
1.2 The Scalar View of Learning . . . . .	3
1.3 Trajectories and Trajectory Manifolds . . . . .	4
<b>2 The Algebra of Repair</b>	<b>6</b>
2.1 Repair Operations and the Relational Character of Persistence . . . . .	6
2.2 The Persistence Invariant and Fixed Points . . . . .	8
2.3 The Repair Hierarchy . . . . .	9
2.4 Repair Triage . . . . .	10
<b>3 Repair as Experiment</b>	<b>12</b>
3.1 The Epistemic Status of Failed Repairs . . . . .	12
3.2 The Memory Requirement . . . . .	13
3.3 The Admissibility Log . . . . .	15
3.4 Repair Triage as the Primary Operation of Learning . . . . .	16

<b>II</b>	<b>Persistence Theory</b>	<b>17</b>
<b>4</b>	<b>Why Some Anomalies Survive</b>	<b>18</b>
4.1	The Information Content of Persistence . . . . .	18
4.2	Parameter Error versus Structural Error . . . . .	18
4.3	The Spectral Decomposition of Persistence . . . . .	19
4.4	Repair Curvature and Repair Entropy . . . . .	21
4.5	Local and Global Persistence; Fragility . . . . .	21
<b>5</b>	<b>The Anatomy of Stubbornness</b>	<b>23</b>
5.1	Historical Cases . . . . .	23
5.1.1	The Perihelion of Mercury . . . . .	23
5.1.2	The Michelson–Morley Experiment . . . . .	24
5.1.3	Blackbody Radiation and the Ultraviolet Catastrophe . . . . .	24
5.2	Common Structure . . . . .	25
<b>III</b>	<b>Topological and Categorical Obstructions</b>	<b>26</b>
<b>6</b>	<b>A Hierarchy of Discrepancy</b>	<b>27</b>
6.1	Two Orthogonal Axes . . . . .	27
6.1.1	Axis I: Type of Discrepancy . . . . .	27
6.1.2	Axis II: Persistence Character . . . . .	28
6.2	The Discrepancy Matrix . . . . .	29
6.3	Why Types Matter . . . . .	29
<b>7</b>	<b>Non-Abelian Tears</b>	<b>31</b>
7.1	Abelian and Non-Abelian Structure . . . . .	31
7.2	Small Metric Distance, Large Topological Distance . . . . .	32

7.3	Holonomy as a Diagnostic . . . . .	33
7.4	Saliency as Accumulated Curvature . . . . .	35
<b>8</b>	<b>The Persistence–Obstruction Schema</b>	<b>36</b>
8.1	Statement . . . . .	36
8.2	The Convergence Gap . . . . .	38
8.3	The Repair Landscape . . . . .	39
8.4	The Persistence Principle . . . . .	40
<b>IV</b>	<b>Admissibility and Ontology Boundaries</b>	<b>41</b>
<b>9</b>	<b>Admissibility Manifolds</b>	<b>42</b>
9.1	Admissibility as Continuation . . . . .	42
9.2	The Boundary of Admissibility . . . . .	43
9.3	RSVP Connections . . . . .	43
<b>10</b>	<b>Ontology Revision</b>	<b>45</b>
10.1	When the Coordinate Chart is Exhausted . . . . .	45
10.2	Representational Inertia and the Rarity of Revision . . . . .	45
10.3	The Dissolving Tear . . . . .	47
10.4	Ontological Saliency and Representational Divergence . . . . .	48
<b>V</b>	<b>Learning and Representation</b>	<b>50</b>
<b>11</b>	<b>Against Error Minimization</b>	<b>51</b>
11.1	The Standard Account and Its Concealed Limits	51
11.2	What Error Minimization Cannot See . . . . .	52

---

11.3 Gradient Descent as Level-0 Repair . . . . .	52
<b>12 Persistence-Driven Learning</b>	<b>54</b>
12.1 Tracking the Repair History . . . . .	54
12.2 Persistence-Weighted Attention . . . . .	55
12.3 Detecting Representational Boundaries During Training . . . . .	55
<b>13 Ontology-Expanding Learning</b>	<b>57</b>
13.1 Learning as Representational Growth . . . . .	57
13.2 Persistent Anomalies as Generators . . . . .	58
13.3 The Geometry of Repair Space . . . . .	58
13.4 CLIO Projection Failures as Ontology Revision	59
<b>VI Suppressed Tears</b>	<b>61</b>
<b>14 How Tears Get Hidden</b>	<b>62</b>
14.1 The Dual of the Persistence Principle . . . . .	62
14.2 Three Mechanisms of Suppression . . . . .	63
14.2.1 Expressive Algebra Containment . . . . .	63
14.2.2 Commitment Asymmetry . . . . .	63
14.2.3 Promotion Blindness . . . . .	64
<b>15 Epicycles and the Convergence Trap</b>	<b>65</b>
15.1 The Ptolemaic System as a Case Study . . . . .	65
15.2 The Convergence Trap . . . . .	66
<b>16 Overfitting and the Ptolemaic Principle</b>	<b>67</b>
16.1 Classical Overfitting and Its Topological Rein- terpretation . . . . .	67
16.2 Generalization as Topological Compatibility . . . . .	68

---

16.3	The Ptolemaic Principle . . . . .	69
<b>VII</b>	<b>Applications</b>	<b>70</b>
<b>17</b>	<b>Mercury and General Relativity</b>	<b>71</b>
17.1	The Anomaly . . . . .	71
17.1.1	Level-0 Repairs . . . . .	71
17.1.2	Level-1 Repairs . . . . .	71
17.1.3	Level-2 Repairs . . . . .	72
17.1.4	The Ontology Revision . . . . .	72
17.2	What the Persistence Framework Adds . . . . .	72
<b>18</b>	<b>Scientific Revolutions as Ontology Revision</b>	<b>74</b>
18.1	A Persistence-Based Account of Scientific Change	74
18.2	The Role of Anomaly Persistence in Discovery	75
18.3	Science as Collective Triage . . . . .	75
<b>19</b>	<b>Biological Adaptation</b>	<b>77</b>
19.1	Evolution as Persistent Anomaly Processing . . . . .	77
19.2	Speciation as Ontology Revision . . . . .	78
19.3	Developmental Regulation . . . . .	78
<b>20</b>	<b>Machine Learning and Representation</b>	<b>80</b>
20.1	Foundation Models and the Persistence Problem	80
20.2	Representation Learning and Topological Fidelity . . . . .	81
20.3	A Research Agenda . . . . .	82
<b>21</b>	<b>RSVP, CLIO, and Admissibility Fields</b>	<b>84</b>
21.1	The RSVP Framework . . . . .	84
21.2	Non-Abelian Tears in RSVP . . . . .	85

---

21.3 CLIO Projection and Representational Com- pression . . . . .	85
21.4 Lamphrodyne Relaxation and Tear Dissolution	86
<b>Conclusion</b>	<b>88</b>
<b>Bibliography</b>	<b>92</b>

## Preface

This monograph began with a simple observation about Mercury.

In the nineteenth century, astronomers knew that the perihelion of Mercury's orbit precessed at a rate that Newtonian mechanics could not fully account for. The discrepancy was small—a matter of forty-three arcseconds per century—and yet it survived every repair that Newtonian astronomy could offer. New planetary bodies were hypothesized and searched for. Atmospheric drag was measured and found insufficient. The oblate shape of the Sun was estimated and incorporated. None of it worked. The anomaly persisted [44, 8].

The standard narrative presents this as a story about a wrong theory waiting to be replaced [24, 25]. But that framing obscures the more important point: the anomaly was not valuable because it was large. It was valuable because it was *stubborn*. Its persistence under an expanding sequence of repair operations was itself information—information not about Mercury's orbit, but about the geometric structure of the space in which the theory was trying to represent that orbit.

This observation motivates the central thesis of the present work.

**Central Thesis.** *Persistent anomalies are evidence of missing representational structure, not merely inaccurate parameters.*

Everything that follows is a development and mathematization of this idea. We introduce a general theory of repair algebras, define a hierarchy of discrepancy types, develop a Persistence–Obstruction Schema connecting repair-invariant anomalies to obstructions in the representational geometry, and propose a framework for ontology-expanding learning in which persistent anomalies function as generators of representational growth rather than targets of error minimization.

Throughout, we maintain the distinction between two modes of failure: the failure of a model to fit observed data, and the failure of a model’s representational architecture to contain any trajectory that could fit observed data. The first mode is addressed by learning. The second mode requires ontology revision. The monograph is, at its core, a mathematical account of how to distinguish them—and of why the distinction so often goes undetected.

Several conceptual innovations introduced in the course of the argument deserve special mention at the outset. Persistence is shown to be relational: a property not of an anomaly alone but of the pair  $(\varepsilon, \mathcal{R})$  where  $\mathcal{R}$  is the repair algebra available to the observer. Memory is shown to be not merely useful but constitutively necessary for the computation of persistence: without a remembered repair trajectory, no anomaly can ever be recognized as persistent, and the entire hierarchy of ontology revision collapses. The geometry of repair space itself—the landscape of available transformations acting on the model manifold—is identified as the ultimate object of

study.

The resulting picture synthesizes themes from philosophy of science, information geometry, algebraic topology, and adaptive systems [24, 1, 19, 11]. The RSVP framework of relational scalar-vector-potential cosmology appears in Part VII as one application domain among several, alongside modern machine learning and scientific epistemology.

*Flyxion, Canada, 2026*

Part I

Repair Theory

# CHAPTER 1

## Models, Observations, and Residuals

The most important question in science is not:  
what is the answer? It is: what kind of question  
is this?

---

*attributed to various sources*

### 1.1 Basic Setup

We begin with the most abstract possible setting for the relationship between a model and its observations.

**Definition 1.1** (Model). A *model* is a triple  $(\mathcal{M}, \Sigma, \pi)$  where  $\mathcal{M}$  is a state space,  $\Sigma$  is a space of observable outputs, and  $\pi : \mathcal{M} \rightarrow \Sigma$  is a prediction map.

A model generates predictions. Observations are elements of  $\Sigma$  that are not necessarily in the image of  $\pi$ . The discrepancy between predictions and observations is what we call a residual.

**Definition 1.2** (Residual). Given a prediction  $\hat{x} = \pi(m)$  for

some  $m \in \mathcal{M}$  and an observation  $x_{\text{obs}} \in \Sigma$ , the *residual* is

$$\varepsilon = x_{\text{obs}} - \hat{x}$$

when  $\Sigma$  is a normed space, or more generally a section of an appropriate bundle over  $\Sigma$  measuring the discrepancy between prediction and observation.

We deliberately leave the notion of “discrepancy” flexible at this stage. In later chapters it will acquire geometric and topological content. For now it suffices to have some measure of how far the prediction misses the observation.

## 1.2 The Scalar View of Learning

Classical learning theory reduces the quality of a model to a single number at each point in observation space [42, 6, 39]. A model with small expected loss is considered adequate.

**Definition 1.3** (Loss-Minimizing Learner). A *loss-minimizing learner* selects

$$m^* = \arg \min_{m \in \mathcal{M}} \mathbb{E}[L(\pi(m), x_{\text{obs}})]$$

for some fixed loss function  $L : \Sigma \times \Sigma \rightarrow \mathbb{R}_{\geq 0}$ .

This scalar view is not wrong. For a vast range of practical problems it produces excellent models. The critique developed in this monograph is that it succeeds in a way that systematically conceals its own limits. The loss function cannot ask whether the representational architecture is adequate [45, 40]; it can only ask how well the current architecture performs. We will argue throughout this monograph

that the reduction of model quality to a scalar at each observation is not merely a computational convenience but an ontological commitment—one that systematically suppresses certain categories of information about model quality.

### 1.3 Trajectories and Trajectory Manifolds

Many interesting observations arise not as isolated points but as trajectories: sequences or paths through observation space indexed by time or some other parameter.

**Definition 1.4** (Trajectory Manifolds). Let  $\gamma_p, \gamma_o : [0, T] \rightarrow \Sigma$  denote the predicted and observed trajectories respectively. As initial conditions vary, these sweep out submanifolds  $M_{\text{pred}}$  and  $M_{\text{obs}}$  of the trajectory space  $\Gamma(\Sigma)$ . We call these the *predicted* and *observed trajectory manifolds*.

The relationship between  $M_{\text{pred}}$  and  $M_{\text{obs}}$ —metric, geometric, topological, and categorical—is the central object of study in this monograph. A pointwise residual  $\|\gamma_o(t) - \gamma_p(t)\|$  discards most of the interesting information carried by this pair. The tools developed below allow us to recover it.

---

#### Derivation 1: Residuals and Repair Sequences

Let  $M$  be a model,  $o_t$  an observation, and  $\hat{o}_t = M(h_t)$  the prediction generated from history  $h_t$ . The ordinary residual is

$$\varepsilon_t = o_t - \hat{o}_t.$$

A repair operation is a map  $R_i : \mathcal{E} \rightarrow \mathcal{E}$  acting on residual

---

space  $\mathcal{E}$ . A sequence of repairs produces

$$\varepsilon_t^{(n)} = R_n R_{n-1} \cdots R_1(\varepsilon_t).$$

The anomaly is *transient* if  $\lim_{n \rightarrow \infty} |\varepsilon_t^{(n)}| = 0$ , and *persistent* if  $\liminf_{n \rightarrow \infty} |\varepsilon_t^{(n)}| > 0$ . Persistence is therefore not simply error magnitude. It is the failure of error to vanish under admissible repair.

---

## CHAPTER 2

# The Algebra of Repair

### 2.1 Repair Operations and the Relational Character of Persistence

**Definition 2.1** (Repair Operation). A *repair operation* is a map  $R : \mathcal{E} \rightarrow \mathcal{E}$  on the space of residuals satisfying  $\|R(\varepsilon)\| \leq \|\varepsilon\|$  for all  $\varepsilon$ , with  $R(\varepsilon) = 0$  whenever  $\varepsilon$  is attributable to a specific class of modifiable parameters within the model.

Repair operations model the actions available to a learning system: adjusting parameters, changing coordinate systems, adding mechanisms, revising priors. Different repair operations correspond to different levels of intervention in the model.

**Example 2.2.** For a physical model with measurement uncertainty  $\sigma$ , a Level-0 repair operation might subtract estimated bias:

$$R_0(\varepsilon) = \varepsilon - \hat{\mu}$$

where  $\hat{\mu}$  is the estimated systematic error. This repair is admissible whenever the model has a free parameter corresponding to measurement offset.

**Definition 2.3** (Repair Algebra). A *repair algebra*  $\mathcal{R} = \{R_i\}_{i \in I}$  is a collection of repair operations closed under composition. The algebra is *admissible* for a model  $\mathcal{M}$  if every  $R_i \in \mathcal{R}$  corresponds to a modification available within  $\mathcal{M}$ 's representational architecture.

The admissibility condition is crucial. A repair algebra for Newtonian celestial mechanics includes operations like adjusting masses, adding new gravitational bodies, and refining orbital parameters. It does not include changing the background geometry from flat to pseudo-Riemannian space—that operation lies outside the algebra. The notion that scientific progress proceeds through structured modification rather than wholesale replacement has deep precedents in research programme methodology [25].

This observation leads immediately to the relational character of persistence. The same anomaly can be transient relative to one repair algebra and persistent relative to another. A modern neural network can absorb anomalies that would immediately force promotion in a simpler model. A Newtonian astronomer and an Einsteinian physicist have access to different repair algebras, and the same residual presents very differently to each.

---

### Derivation 2: Persistence Relative to a Repair Algebra

Let  $\mathcal{R}$  be the repair algebra generated by admissible operations  $R_i$ . Define the *persistence* of an anomaly as

$$P(\varepsilon, \mathcal{R}) = \inf_{R \in \langle \mathcal{R} \rangle} \|R(\varepsilon)\|.$$

If  $P(\varepsilon, \mathcal{R}) = 0$ , the anomaly is repairable within the current ontology. If  $P(\varepsilon, \mathcal{R}) > 0$ , the anomaly contains structure not removable by the available repair algebra. This makes persistence relational: the same anomaly may be transient for one observer and persistent for another, depending on what repairs are available.

---

## 2.2 The Persistence Invariant and Fixed Points

**Definition 2.4** (Repair Closure and Persistence Invariant). Define the *repair closure operator*

$$\mathcal{C}_{\mathcal{R}}(\varepsilon) = \lim_{n \rightarrow \infty} R_n \cdots R_1(\varepsilon),$$

where the limit is taken over optimal repair sequences. The *persistence invariant* is

$$\text{Pers}_{\mathcal{R}}(\varepsilon_0) = \|\mathcal{C}_{\mathcal{R}}(\varepsilon_0)\|.$$

An anomaly is *transient* if  $\text{Pers}_{\mathcal{R}}(\varepsilon_0) = 0$  and *persistent* if  $\text{Pers}_{\mathcal{R}}(\varepsilon_0) > 0$ .

The central claim of this monograph is that persistence carries more information than magnitude. A residual of magnitude  $10^{-3}$  that is persistent is more informative about representational limits than a residual of magnitude  $10^3$  that is transient.

When the persistence invariant is non-zero, the limit  $\varepsilon^* = \mathcal{C}_{\mathcal{R}}(\varepsilon_0)$  satisfies a fixed-point condition.

**Proposition 2.5** (Persistence Fixed Point). *Let  $\mathcal{R}$  be a contractive repair algebra on a complete residual space  $\mathcal{E}$ . Then*

every persistent anomaly contains a non-zero fixed-point component  $\varepsilon^*$  satisfying

$$R(\varepsilon^*) = \varepsilon^* \quad \text{for all admissible } R \in \mathcal{R}.$$

*Proof.* By contractivity, the sequence  $R_n \cdots R_1(\varepsilon_0)$  converges in  $\mathcal{E}$  by the Banach fixed-point theorem. If the limit  $\varepsilon^*$  is non-zero, then applying any  $R \in \mathcal{R}$  to  $\varepsilon^*$  produces a further contraction. Since  $\varepsilon^*$  is already the infimum of the repair sequence,  $\|R(\varepsilon^*)\| \geq \|\varepsilon^*\|$ . Combined with  $\|R(\varepsilon^*)\| \leq \|\varepsilon^*\|$ , we obtain  $R(\varepsilon^*) = \varepsilon^*$ .  $\square$

This gives a precise interpretation: *persistent anomalies are fixed points of repair*. They are the components of a residual that the repair dynamics cannot move, the irreducible remainder after all available repair has been exhausted.

### 2.3 The Repair Hierarchy

Real learning systems do not apply all repair operations simultaneously. They apply them in order of cost, preference, or availability. This gives rise to a natural hierarchy.

**Definition 2.6** (Repair Hierarchy). A *repair hierarchy* is a sequence of repair algebras

$$\mathcal{R}_0 \subset \mathcal{R}_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_K$$

where each  $\mathcal{R}_k$  is a strict extension of  $\mathcal{R}_{k-1}$  with higher operational cost.

Level-0 repairs are cheapest: spelling corrections, parameter nudges, calibration adjustments. Level-1 repairs are more

expensive: coordinate changes, mechanism additions. Higher levels involve more fundamental alterations to the model's architecture.

**Definition 2.7** (Promotion). An anomaly  $\varepsilon$  is *promoted* from level  $k$  to level  $k + 1$  when

$$\text{Pers}_{\mathcal{R}_k}(\varepsilon) > 0 \quad \text{but the system attempts } R \in \mathcal{R}_{k+1}.$$

Promotion occurs when every cheaper repair has been exhausted.

The promotion process gives a procedural definition of a tear: an anomaly that has survived every repair at every level below some threshold is a candidate for a structural, rather than parametric, diagnosis.

## 2.4 Repair Triage

**Definition 2.8** (Repair Triage). *Repair triage* is the process by which a learning system:

- (i) presents an anomaly to the lowest available repair level,
- (ii) promotes it if the repair fails,
- (iii) continues promotion until either the anomaly is absorbed or the highest available level is reached,
- (iv) flags anomalies that survive all levels as candidates for ontology revision.

Under this view, learning is not primarily error reduction. It is the continuous operation of a triage process that sorts

---

anomalies by their structural depth. Most anomalies are resolved at low levels. The valuable ones are those that climb.

## CHAPTER 3

# Repair as Experiment

### 3.1 The Epistemic Status of Failed Repairs

Each repair attempt is an experiment: a test of the hypothesis that the anomaly is attributable to a specific class of correctable features [35, 17]. When a repair fails—when the anomaly survives—the experiment has a negative result. As repair attempts accumulate, the surviving anomaly is progressively purified. Each failed repair eliminates one more explanation that does not work. What remains after exhaustive application of a repair algebra is, in a precise sense, the irreducible content of the anomaly.

**Proposition 3.1** (Purification by Repair). *Let  $\varepsilon_0$  be an anomaly and let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be a finite repair algebra. Then  $\text{Pers}_{\mathcal{R}}(\varepsilon_0)$  represents the component of  $\varepsilon_0$  that is orthogonal to every repair direction in  $\mathcal{R}$ .*

*Sketch.* Each  $R_i$  defines a subspace of the residual space in which it can reduce anomalies. The image of the full repair sequence is the projection onto the orthogonal complement of the span of these subspaces. The persistence invariant is the norm of this projection.  $\square$

The repair sequence is therefore not merely a sequence of attempts. It is a measurement process. The outcome of the measurement is the persistence invariant. A sequence that converges to zero reports: this anomaly was local. A sequence that plateaus reports: this anomaly has revealed a boundary.

### 3.2 The Memory Requirement

A critical but easily overlooked requirement for repair triage is memory. A system without memory of its own repair history cannot recognize that an anomaly is old. It treats each repair attempt as fresh, potentially repeating the same repair indefinitely while the anomaly remains persistent. The survival record does not exist, and therefore promotion cannot occur.

**Theorem 3.2** (No Persistence Without Memory). *Let  $L$  be a memoryless learner seeing only the current residual  $\varepsilon_t$ . Then there exists no function  $f(\varepsilon_t)$  capable of computing  $P(\varepsilon, \mathcal{R})$ . Persistence necessarily depends on a trajectory  $(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_n)$ . Memory is therefore a necessary condition for persistence detection.*

*Proof.* Suppose for contradiction that  $P(\varepsilon, \mathcal{R}) = f(\varepsilon_t)$  for some function  $f$ . Consider two repair histories: (A) anomaly  $\varepsilon$  encountered for the first time, with no prior repairs applied; and (B) anomaly  $\varepsilon$  encountered after  $n$  failed repairs in diverse repair classes. In both cases  $\varepsilon_t = \varepsilon$ , so  $f(\varepsilon_t)$  takes the same value. But  $P(\varepsilon, \mathcal{R})$  in case (B) is strictly greater: in (B) we know  $n$  repair classes have been exhausted, establishing

a lower bound on persistence absent in (A). Therefore  $f(\varepsilon_t)$  cannot compute  $P(\varepsilon, \mathcal{R})$  from  $\varepsilon_t$  alone.  $\square$

This demotes memory from a supporting concept to a foundational one. Earlier frameworks might have said: memory stores the history of learning. The repair-triage framework implies something stronger. This view aligns with broader perspectives that treat memory as constitutive of adaptive intelligence rather than passive storage [21, 11].

---

### **Derivation 6: Memory as the Condition for Persistence**

Persistence cannot be computed from a single residual. It requires a remembered repair trajectory:

$$\varepsilon \rightarrow R_1(\varepsilon) \rightarrow R_2R_1(\varepsilon) \rightarrow \cdots \rightarrow R_n \cdots R_1(\varepsilon).$$

A memoryless learner sees only  $\varepsilon_t$ . A memory-bearing learner sees the sequence  $\{\varepsilon^{(0)}, \varepsilon^{(1)}, \dots, \varepsilon^{(n)}\}$ . Memory is therefore not merely storage. It is the structure that makes persistence measurable. The dependency chain follows:

Memory  $\Rightarrow$  Persistence  $\Rightarrow$  Promotion  $\Rightarrow$  Ontology Revision  $\Rightarrow$  Learning

Most theories treat memory as a component of learning. The repair-triage framework suggests that learning itself may be a consequence of maintaining persistent repair histories long enough for anomalies to climb the hierarchy.

---

### 3.3 The Admissibility Log

The operational implementation of repair triage requires a structured record. We call this the *Admissibility Log*: a data structure that stores not merely whether a repair failed, but the full repair path.

**Definition 3.3** (Admissibility Log). An *Admissibility Log* for an anomaly  $\varepsilon_0$  is a record

$$(\varepsilon_0, R_1, \varepsilon_1, R_2, \varepsilon_2, \dots, R_n, \varepsilon_n)$$

storing the sequence of repairs attempted and residuals produced. The Log enables computation of the persistence trajectory and supports detection of path-dependence in the repair sequence.

**Definition 3.4** (Admissibility Pressure). Define the *admissibility pressure* accumulated after  $n$  repairs as

$$A_n = \sum_{i=1}^n w_i \cdot \mathbf{1}(R_i \text{ failed}),$$

where  $w_i$  weights repairs by level:  $w_i = k$  for a level- $k$  repair. Large  $A_n$  indicates repeated failed repairs across multiple levels.

**Proposition 3.5** (Promotion Threshold). *If admissibility pressure exceeds a critical value  $A_n > A_c$ , then every rational repair policy promotes the anomaly to a higher repair level, since the expected cost of further low-level repair exceeds the expected benefit of persistence reduction.*

*Sketch.* The expected utility of continuing level- $k$  repair after  $n$  failures is  $B_k \cdot \text{Pr}(\text{success}) - C_k$ , where  $B_k$  is the benefit of reducing persistence at level  $k$  and  $C_k$  is the cost per attempt. After  $n$  failures, Bayesian updating drives  $\text{Pr}(\text{success}) \rightarrow 0$ . For  $A_c$  corresponding to the point at which this expectation becomes negative, rational policy promotes.  $\square$

The Admissibility Log supports four diagnostic signals: the current residual size  $|\varepsilon_n|$ ; the residual slope  $d|\varepsilon_n|/dn$  measuring how quickly repairs are becoming less effective; the repair diversity, counting how many distinct repair classes have been attempted; and the loop inconsistency, measuring whether different repair paths from the same starting point converge to equivalent endpoints.

### 3.4 Repair Triage as the Primary Operation of Learning

We may now offer a recharacterization of learning that will be developed throughout the rest of this monograph. Under the repair-triage view, learning is not primarily error reduction. It is the continuous operation of a triage process that sorts anomalies by their structural depth. Most anomalies are resolved at low levels. The valuable ones are those that climb.

# Part II

## Persistence Theory

## CHAPTER 4

# Why Some Anomalies Survive

### 4.1 The Information Content of Persistence

The question we must now address is not merely empirical—not simply “some anomalies happen to be persistent”—but explanatory: *why* do some anomalies survive every repair? What structure in the world, or in the relationship between the world and the model, causes them to be immune to repair?

We propose the following answer, which the remainder of Part II and all of Part III will develop rigorously.

An anomaly is persistent because it encodes information about the *shape of the boundary* of the model’s representational space. It is not pointing at a fact about the object being modeled. It is pointing at a fact about the coordinate system being used to model it. The repair algebra cannot remove it because no operation in the algebra addresses coordinates—only contents.

### 4.2 Parameter Error versus Structural Error

**Definition 4.1** (Parameter and Structural Error). A residual  $\varepsilon$  is a *parameter error* if  $x_{\text{obs}} \in \pi(\mathcal{M})$ —the observation lies in

the model's image and the model simply has wrong parameter values. It is a *structural error* if  $x_{\text{obs}} \notin \pi(\mathcal{M})$ : no parameter adjustment can remove it.

The distinction parallels that between parameter estimation and model misspecification in statistical learning theory [42, 28].

**Proposition 4.2** (Persistence Detects Structural Errors). *Let  $\mathcal{R}_{\text{param}}$  be the repair algebra of all parameter adjustments within  $\mathcal{M}$ . If  $\text{Pers}_{\mathcal{R}_{\text{param}}}(\varepsilon) > 0$ , then  $\varepsilon$  is a structural error.*

*Proof.* If  $\varepsilon$  is a parameter error, there exists  $m' \in \mathcal{M}$  with  $\pi(m') = x_{\text{obs}}$ , so the adjustment  $R : m \mapsto m'$  satisfies  $R(\varepsilon) = 0$ . Contrapositive gives the result.  $\square$

A parameter error is always transient: the repair operation of adjusting parameters can reduce it to zero. A structural error may be persistent even under infinitely many parameter adjustments, because parameter adjustment does not expand the image of  $\pi$ . The Persistence Principle, in its most basic form, asserts that persistence is a diagnostic for structural error.

### 4.3 The Spectral Decomposition of Persistence

The persistence invariant  $P(\varepsilon, \mathcal{R})$  treats persistence as a scalar. But repairs act differently on different modes of an anomaly, and the richer structure of this action is itself informative. The decomposition of anomalies into repair-sensitive and repair-invariant modes is inspired by spectral methods throughout information geometry and dynamical systems [1, 16].

**Definition 4.3** (Persistence Spectrum). Let  $\mathcal{R}$  act on residual space  $\mathcal{E}$ . Decompose  $\varepsilon = \sum_i \varepsilon_i$  into modes that diagonalize the action of  $\mathcal{R}$ , with *persistence eigenvalues*  $\lambda_i$  defined by

$$R(\varepsilon_i) = \lambda_i \varepsilon_i.$$

The *persistence spectrum* of  $\varepsilon$  is the multiset  $\{\lambda_i\}$ .

Three regimes follow. Modes with  $\lambda_i < 1$  decay under repair: these are the correctable components of the anomaly. Modes with  $\lambda_i = 1$  are invariant: repairs act on them trivially, and they persist regardless of how many repair steps are taken. Modes with  $\lambda_i > 1$  are amplified by repair—the anomaly in these modes becomes more visible as repairs accumulate.

The amplified case is particularly striking. It describes anomalies that repair makes sharper rather than smaller. The historical career of Mercury’s perihelion may have this character: as Newtonian mechanics became more precise and other perturbations were calculated more accurately, the residual of forty-three arcseconds became more undeniable rather than fading into observational noise.

The persistence invariant is then the norm of the invariant subspace:

$$\text{Pers}_{\mathcal{R}}(\varepsilon) = \left\| \sum_{\lambda_i=1} \varepsilon_i \right\|.$$

#### 4.4 Repair Curvature and Repair Entropy

**Definition 4.4** (Repair Curvature and Repair Entropy). Let  $p_n = \|\varepsilon^{(n)}\|$ . The *repair curvature* is

$$\kappa_R = \frac{d^2 p_n}{dn^2}.$$

Positive  $\kappa_R$  means repairs are becoming less effective: the rate of improvement is decelerating, indicating that the anomaly is approaching its persistent component. Let  $S_R(n)$  be the cumulative complexity added by  $n$  repair steps. The *repair efficiency* is

$$E_R = \frac{-\Delta p_n}{\Delta S_R}.$$

A healthy repair sequence has high  $E_R$ . A convergence-trap sequence has  $E_R \rightarrow 0$  as  $S_R$  grows without bound.

Positive repair curvature is a pre-failure signal: it allows a learner to detect the approach to a representational boundary before the residual fully plateaus, providing earlier warning than persistence monitoring alone.

#### 4.5 Local and Global Persistence; Fragility

**Definition 4.5** (Local and Global Persistence). An anomaly has *local persistence* if it survives every local repair—every modification to a bounded region of the model’s parameter space—but disappears under a global coordinate transformation. It has *global persistence* if it survives global transformations as well.

Mercury’s precession is a case of global persistence: it sur-

vived not only local parameter adjustments but every global transformation available within Newtonian mechanics. Distinguishing local from global persistence helps identify the scope of the required repair.

**Definition 4.6** (Fragility). The *fragility* of an anomaly  $\varepsilon$  is

$$F(\varepsilon) = \mathbb{E}_\delta [\|\varepsilon - \delta(\varepsilon)\|]$$

where  $\delta$  ranges over small perturbations of the observation. A genuine structural anomaly should be simultaneously highly persistent and low-fragility: it should be robust to small changes in the observation while resisting repair.

Fragility provides a filter against false positives in persistence detection. The joint criterion  $P(\varepsilon, \mathcal{R}) > 0$  and  $F(\varepsilon)$  small distinguishes genuine structural anomalies from coincidental persistence. A noisy residual may appear persistent under a specific repair sequence while being highly fragile; a genuine tear is stable under both the repair sequence and perturbation of the observation.

## CHAPTER 5

# The Anatomy of Stubbornness

### 5.1 Historical Cases

Before proceeding to the mathematical framework, we examine three historical cases in which small persistent anomalies ultimately required ontology revision. These cases will serve as recurring touchstones throughout the monograph.

#### *5.1.1 The Perihelion of Mercury*

Mercury's perihelion precesses at approximately 574 arcseconds per century. Of this, Newtonian mechanics can account for about 531 arcseconds due to perturbations from other planets. The residual of 43 arcseconds per century survived every repair available within classical mechanics [44]. The anomaly's persistence spectrum contained invariant modes that no Newtonian repair could address, because the invariant subspace lay outside the image of the Newtonian repair algebra. As documented in detail in Part VII, this anomaly climbed every level of the Newtonian repair hierarchy before requiring the replacement of Euclidean spacetime with pseudo-Riemannian geometry.

### 5.1.2 *The Michelson–Morley Experiment*

The 1887 Michelson–Morley experiment [29] found no detectable variation in the speed of light depending on the Earth’s orientation relative to the hypothetical luminiferous ether. The expected signal was small but well within the experiment’s sensitivity. The null result was anomalous: classical mechanics predicted a positive result. Decades of repair attempts—modified ether drag hypotheses, Lorentz–FitzGerald contraction, revised ether models—failed to absorb the anomaly while preserving the broader classical framework. The anomaly’s  $\lambda = 1$  modes corresponded to the invariance of  $c$ , accommodable only by replacing the Galilean transformation group with the Lorentz group.

### 5.1.3 *Blackbody Radiation and the Ultraviolet Catastrophe*

Classical thermodynamics and electromagnetic theory, combined via the equipartition theorem, predicted that a blackbody should radiate infinite energy at high frequencies. The observed spectrum was dramatically different at short wavelengths [33]. Despite extensive attempts to modify the classical derivation, no repair within the classical framework reproduced the observed Planck distribution. The repair entropy of classical attempts was large and growing before Planck introduced energy quantization. This new element—the quantum of action—had no analogue in the classical repair algebra. The repair sequence was branching, not converging, before the ontology revision occurred.

## 5.2 Common Structure

These three cases share a common structure that motivates the theoretical framework developed in subsequent chapters.

In each case:

- (i) The anomaly was *numerically small* relative to the total predictive success of the prevailing theory.
- (ii) The anomaly was *persistent*: it survived an extended sequence of increasingly sophisticated repair attempts within the existing representational framework.
- (iii) *Repair curvature* was positive throughout the failure period, signaling the approach to a boundary before the residual fully plateaued.
- (iv) The eventual resolution required *ontology revision*: the introduction of representational structures absent from the original framework.
- (v) In retrospect, the anomaly was not *about* the phenomenon being modeled (Mercury's orbit, light propagation, thermal radiation). It was *about the geometry of the theory being used to model it*.

This common structure is what the Persistence–Obstruction Schema in Part III will formalize.

## Part III

# Topological and Categorical Obstructions

## CHAPTER 6

# A Hierarchy of Discrepancy

### 6.1 Two Orthogonal Axes

The space of possible anomalies is structured along two orthogonal axes. The first axis concerns *where* the discrepancy lives—its type in a mathematical sense. The second axis concerns *how it behaves under repair*—its persistence character. Together they form a matrix that is considerably richer than any single hierarchy.

#### 6.1.1 Axis I: Type of Discrepancy

**Level 0: Numerical.** The discrepancy is a scalar or vector in a normed space:

$$\delta_0 = \|x_{\text{obs}} - x_{\text{pred}}\|.$$

**Level 1: Geometric.** The discrepancy involves distances between manifolds, Hausdorff metrics, or Riemannian curvature differences:

$$\delta_1 = d_{\text{geom}}(M_{\text{obs}}, M_{\text{pred}}).$$

**Level 2: Topological.** The discrepancy is captured by an obstruction class in a cohomology theory:

$$\delta_2 = [\tau] \in H^k(M; G)$$

for some coefficient group  $G$ . The use of homotopy and cohomology as indicators of structural distinction follows standard treatments in algebraic topology [19, 4].

**Level 3: Categorical.** The discrepancy is the failure of a diagram of morphisms to commute:

$$f \circ g \neq h \circ k$$

even though the model requires  $f \circ g = h \circ k$ . This is a failure at the level of representational architecture itself. Such failures are naturally expressed within category-theoretic frameworks [27, 2, 37].

### 6.1.2 Axis II: Persistence Character

**Transient.** The anomaly vanishes under admissible repair.

**Stable.** The anomaly is reduced but not eliminated; it remains at the same discrepancy type.

**Persistent.** The anomaly survives arbitrarily long repair sequences.

**Invariant.** The anomaly is invariant under the full admissible repair algebra and represents a genuine structural feature.

## 6.2 The Discrepancy Matrix

These two axes organize the space of anomalies as follows.

	<b>Transient</b>	<b>Stable</b>	<b>Persistent</b>
<b>Numerical</b>	Calibration error	Systematic bias	Structural mismatch
<b>Geometric</b>	Coord. choice	Fit residual	Manifold incompleteness
<b>Topological</b>	Local obstruction	Boundary effect	Genuine tear
<b>Categorical</b>	Model bug	Symmetry break	Diagram failure

The cells in the rightmost column represent anomalies where no admissible repair can help. They are the candidates for ontology revision. The cells in the bottom-right corner represent the most structurally deep: categorical invariants that survive all repair and signal that the category of models being considered is itself the wrong one.

## 6.3 Why Types Matter

The type of a discrepancy determines what kind of intervention is appropriate. Applying a numerical repair to a topological discrepancy does not merely fail—it consumes resources while providing false evidence that the anomaly is being addressed.

This is the Ptolemaic failure mode. The repair type was systematically mismatched to the discrepancy type. Epicycles are numerical and geometric repairs. The heliocentric structure of the solar system is a categorical revision. No accumulation of epicycles could address a categorical discrepancy.

ancy, because the operations involved do not interact with the relevant level.

The discrepancy matrix therefore has diagnostic value. When a repair sequence converges without reaching zero, determining which type of discrepancy has been reached allows the system to identify the appropriate level of intervention.

## CHAPTER 7

# Non-Abelian Tears

### 7.1 Abelian and Non-Abelian Structure

The terminology “non-abelian tear” requires explanation. In algebra, a structure is *abelian* if its operations commute:  $a + b = b + a$ , or more generally  $xy = yx$ . A non-abelian structure is one in which the order of operations matters.

The canonical example is rotation in three dimensions. Rotating a rigid body by ninety degrees around the  $x$ -axis followed by ninety degrees around the  $y$ -axis produces a different orientation than performing these operations in the reverse order. The group of rotations  $\text{SO}(3)$  is non-abelian.

The relevance to anomalies is the following. If two trajectory histories can be continuously identified—if one can be smoothly deformed into the other while preserving all relevant structure—then the order in which their constituent operations are composed does not affect the final result. The identification is abelian in this sense. If the histories *cannot* be continuously identified while preserving their compositional structure, then order matters, and the discrepancy is non-abelian.

**Definition 7.1** (Non-Abelian Tear). Let  $M_{\text{pred}}$  and  $M_{\text{obs}}$  be trajectory manifolds over a common base space  $B$ . A *non-abelian tear* is an obstruction to the existence of a homotopy equivalence

$$\phi : M_{\text{pred}} \xrightarrow{\cong} M_{\text{obs}}$$

compatible with the compositional structure of the trajectory algebras. In favorable geometric settings, it is represented by a non-trivial class

$$[\tau] \in H^1(M; \pi_1(\text{Diff}(M)))$$

or, in the categorical setting, a failure of a natural isomorphism between functors representing the two trajectory histories.

*Remark 7.2.* The term “non-abelian” here refers to the non-commutativity of the path algebra, not to a specific group. A tear can be non-abelian even when the underlying manifold is simply connected, provided the obstruction involves composition of trajectories in an essential way.

## 7.2 Small Metric Distance, Large Topological Distance

The crucial property of non-abelian tears is that they can correspond to arbitrarily small metric distances. Two manifolds can be metrically close—every point in one is near some point in the other—yet topologically distinct in ways that no smooth deformation can remove. The obstruction-theoretic viewpoint follows classical constructions in algebraic topology and sheaf theory [41, 23].

The simplest example is the Möbius band and the annulus. They are locally similar and homotopy equivalent to a circle, but they are not homeomorphic as surfaces: one is non-orientable and the other is orientable. A model working in annulus-coordinates cannot represent trajectories that wind non-trivially around the Möbius band, regardless of how finely it samples the metric distance between them.

This is the mathematical content of the earlier observation that “the learner sees small error; the topology sees different universe.” The metric distance measures how close the two descriptions are as collections of points. The topological invariants measure whether they belong to the same homotopy class. These are genuinely orthogonal quantities.

### 7.3 Holonomy as a Diagnostic

The most computationally tractable probe of topological discrepancy is holonomy. Parallel transport around a closed loop in a curved space does not return a vector to its original orientation. The accumulated rotation is the holonomy of the loop, and it measures the curvature enclosed. The geometric interpretation of path-dependence through holonomy is standard in differential geometry and gauge theory [31, 12].

**Definition 7.3** (Trajectory Holonomy Discrepancy). Let  $\ell$  be a closed loop in trajectory space, and let  $P_{\text{pred}}(\ell)$  and  $P_{\text{obs}}(\ell)$  denote parallel transport around  $\ell$  under the predicted and observed connections respectively. The *holonomy discrepancy* is

$$\text{Hol}(\ell) = P_{\text{obs}}(\ell) \circ P_{\text{pred}}(\ell)^{-1}.$$

If  $\text{Hol}(\ell) = \text{id}$  for every loop  $\ell$ , the two connections are

locally equivalent. If  $\text{Hol}(\ell) \neq \text{id}$  for some  $\ell$ , the discrepancy has topological content that no local adjustment can remove.

**Proposition 7.4** (Holonomy Detects Tears). *If  $\text{Hol}(\ell) \neq \text{id}$  for some loop  $\ell$ , then  $M_{\text{pred}} \not\cong M_{\text{obs}}$  in the homotopy category of manifolds with connection.*

The significance of this result is practical: holonomy can in principle be estimated from trajectory data without computing full cohomology. A learner can probe the holonomy of loops in its trajectory space to detect the presence of structural discrepancies before they become visible as large metric errors.

---

### Derivation 5: Non-Abelian Tears as Holonomy

Let  $\gamma_1, \gamma_2$  be two histories connecting the same apparent endpoints in state space. In an abelian regime, the transport operators satisfy  $T_{\gamma_1} = T_{\gamma_2}$ , so order does not matter. In a non-abelian regime,

$$H(\gamma_1, \gamma_2) = T_{\gamma_1} T_{\gamma_2}^{-1} \neq I.$$

The residual transformation  $H(\gamma_1, \gamma_2)$  is the holonomy of the tear. If  $H = I$ , the discrepancy is locally repairable. If  $H \neq I$ , the two histories cannot be identified without adding structure. The commutator test for practical detection: compare  $R_a R_b(\varepsilon)$  versus  $R_b R_a(\varepsilon)$ . If  $[R_a, R_b](\varepsilon) \neq 0$  across diverse repair families, the anomaly has path-dependence—the operational signature of a non-abelian tear. Persistent error says “something is hard to fix.” Persistent non-commutation says “the representation is missing structure.” The Admissibility

---

Log stores the full repair path precisely to enable this test.

---

## 7.4 Saliency as Accumulated Curvature

**Definition 7.5** (Holonomy Saliency). The *saliency* of a region  $U$  in state space is proportional to the maximum holonomy discrepancy accumulated by loops passing through  $U$ :

$$\text{Sal}(U) \propto \sup_{\ell \ni U} \|\text{Hol}(\ell)\|.$$

Under this definition, saliency is not the magnitude of local prediction error. It is the accumulated curvature of the discrepancy between predicted and observed trajectory bundles. A region can have zero local error and high saliency if the discrepancy has curvature that concentrates around it. This is precisely the regime in which conventional anomaly detection fails: a learner that monitors  $\|x_{\text{obs}} - x_{\text{pred}}\|$  pointwise will assign zero saliency to a region with zero local error, even if that region lies at the center of a non-trivial holonomy loop.

## CHAPTER 8

# The Persistence–Obstruction Schema

### 8.1 Statement

We now state the central result of this monograph. It establishes the bridge between repair-theoretic persistence and topological obstruction: not merely that they are related, but that persistence *is* obstruction, in a precise cohomological sense. The result is stated as a principle and schema rather than a theorem, because its precise form depends on the geometric setting and the conclusion holds in full generality only under assumptions that must be verified case by case [19, 41].

**Principle 8.1** (Persistence–Obstruction Schema). Let  $\mathcal{M}$  be a model with admissible repair algebra  $\mathcal{R}$ , and let  $\varepsilon_0$  be an anomaly with  $\text{Pers}_{\mathcal{R}}(\varepsilon_0) > 0$ . Then there is an obstruction to continuously extending the repair sequence to zero, represented—in favorable geometric settings, particularly when  $\mathcal{M}$  carries a manifold structure and  $\mathcal{R}$  acts by diffeomorphisms—by a non-trivial cohomological class

$$[\tau(\varepsilon_0)] \neq 0$$

in an appropriate cohomology theory associated with  $\mathcal{M}$ , on which every admissible repair acts trivially.

*Constructive argument for the manifold case.* Let  $\{R_n\}$  be an arbitrary admissible repair sequence and let  $\varepsilon_n = R_n \cdots R_1(\varepsilon_0)$ . By hypothesis,  $\inf_n \|\varepsilon_n\| > 0$  for every such sequence. Consequently there exists an open set  $U \subset \mathcal{E}$  containing 0 such that no repair path enters  $U$ . The complement  $\mathcal{E} \setminus U$  is closed and contains all repair paths. The residual  $\varepsilon_0$  lies in a connected component  $C$  of  $\mathcal{E} \setminus U$  that admits no path to 0, defining a non-trivial element in  $H^0(\mathcal{E} \setminus U; \mathbb{Z})$ .

For the cohomological refinement: the repair algebra  $\mathcal{R}$  acts on  $\mathcal{E}$  by a family of maps, defining a sheaf  $\mathcal{S}_{\mathcal{R}}$  over the space of model parameters. Persistence of  $\varepsilon_0$  means the stalk of this sheaf at 0 does not contain the section defined by  $[\varepsilon_0]$ . The obstruction to extending this section to 0 is a class in  $H^1(\mathcal{M}; \mathcal{S}_{\mathcal{R}})$ . This class is non-trivial by assumption, and every admissible repair acts on it by a coboundary, hence trivially in cohomology.  $\square$

*Remark 8.2.* The precise cohomology theory varies by context. For trajectory manifolds, the relevant class is typically in  $H^1(M; \pi_1(\text{Aut}(M)))$ . For categorical models, it is a natural transformation obstruction in  $\text{Ext}^1(\mathbf{C}, \mathbf{D})$ . For learning systems, it corresponds to representational non-identifiability. In all cases, the schema licenses the inference: persistent anomaly implies structural obstruction.

---

### **Derivation 7: The Persistence–Obstruction Principle**

Let  $\varepsilon$  be an anomaly and  $\mathcal{R}$  an admissible repair algebra. If

$P(\varepsilon, \mathcal{R}) > 0$ , then there exists an obstruction class  $[\tau(\varepsilon)] \neq 0$  in the representational geometry associated with the model. Informally:

persistent anomaly  $\Rightarrow$  non-trivial obstruction  $\Rightarrow$  missing structure.

The repair sequence is an experiment. If its limit is zero, the anomaly was local. If its limit is non-zero, the anomaly has revealed a boundary of the model's admissibility manifold. The obstruction class  $[\tau(\varepsilon)]$  characterizes the shape of that boundary.

---

## 8.2 The Convergence Gap

The mathematical heart of the Ptolemaic Principle is a single observation about convergence in two different spaces.

---

### Derivation: The Convergence Gap

Let  $\varepsilon_n \rightarrow 0$  while  $[\tau_n] = [\tau_0] \neq 0$ . Then

$$\lim_{n \rightarrow \infty} \|\varepsilon_n\| = 0 \quad \text{does not imply} \quad \lim_{n \rightarrow \infty} [\tau_n] = 0.$$

Convergence in residual space is strictly weaker than convergence in representational space. A repair sequence can drive the numerical manifestation of an anomaly to zero while leaving its cohomological class intact. The tear has not been dissolved. It has been hidden.

---

### 8.3 The Repair Landscape

**Definition 8.3** (Repair Graph and Repair Landscape). Define the *repair graph*  $G_{\mathcal{R}} = (V, E)$  where vertices are model states and edges are admissible repairs. The *repair distance* between states  $x$  and  $y$  is the shortest path length  $d_{\mathcal{R}}(x, y)$  in this graph. The *repair landscape* of  $\mathcal{M}$  is the metric space  $(\mathcal{M}, d_{\mathcal{R}})$ .

---

#### Derivation 8: The Repair Landscape

The repair graph  $G_{\mathcal{R}}$  makes ontology boundaries visible as disconnected components. When  $d_{\mathcal{R}}(x, y) = \infty$ —when no finite sequence of admissible repairs connects  $x$  to  $y$ —the two states belong to different connected components of repair space. Ontology revision is the operation of crossing a disconnected boundary: acquiring a repair operation that connects previously unreachable components. The repair landscape therefore encodes both the capacity and the limits of the current representational architecture in a single geometric object.

---

Two models may produce identical residuals and identical persistence invariants yet have dramatically different repair landscapes. One may have many short repair paths to the correct ontology; another may require crossing enormous barriers. The landscape determines not just whether ontology revision is possible but how difficult it is to reach.

## 8.4 The Persistence Principle

**Principle 8.4** (Persistence Principle). Let  $\mathcal{M}$  be a model and  $\mathcal{R}$  its admissible repair algebra. An anomaly satisfying  $\text{Pers}_{\mathcal{R}}(\varepsilon) > 0$  encodes information about the representational structure of  $\mathcal{M}$  rather than information about the state being represented:

$$\text{Pers}_{\mathcal{R}}(\varepsilon) > 0 \implies \text{structural obstruction in } (\mathcal{M}, \mathcal{R}).$$

In words: a persistent anomaly is not a fact about the world. It is a fact about the geometry of the model used to represent the world.

## Part IV

# Admissibility and Ontology Boundaries

## CHAPTER 9

# Admissibility Manifolds

### 9.1 Admissibility as Continuation

A model does not merely assign probabilities or predictions to states. It defines a region through which trajectories can be coherently extended. This region is the admissible domain of the model.

**Definition 9.1** (Admissibility Manifold). The *admissibility manifold*  $\mathcal{A}(\mathcal{M})$  of a model  $\mathcal{M}$  is the maximal subspace of  $\Sigma$  through which the model's prediction map  $\pi$  admits a continuous extension under the dynamics generated by  $\mathcal{M}$ .

Admissibility is a property of continuation rather than probability. A state may have very low probability under a model's distribution and still be admissible—the model can track it, explain it, continue the trajectory through it. Conversely, a state may appear likely under some measure but be inadmissible if the model cannot generate a coherent trajectory passing through it.

**Example 9.2.** In classical Newtonian mechanics, the admissibility manifold is essentially all of phase space: any initial

condition generates a unique forward trajectory. But within the context of a specific model of the solar system—with fixed masses and known planets—the admissible region is the set of initial conditions compatible with observed planetary constraints. Mercury’s observed trajectory is admissible in this restricted sense: Newtonian mechanics can track it. What is inadmissible is the specific precession rate, which the model cannot reproduce regardless of parameter choice.

## 9.2 The Boundary of Admissibility

The boundary of  $\mathcal{A}(\mathcal{M})$  is the set of states at which the model’s coordinate chart begins to fail. Near the boundary, trajectories become increasingly difficult to extend: the model requires more and more representational resources to track what is happening, without successfully doing so.

**Definition 9.3** (Ontology Boundary). The *ontology boundary* of a model  $\mathcal{M}$  is the locus of states at which no admissible repair path can restore trajectory continuity.

The ontology boundary is where non-abelian tears live. A trajectory approaching the ontology boundary accumulates holonomy: the discrepancy between predicted and observed parallel transport grows without any admissible repair being able to stop it.

## 9.3 RSVP Connections

The RSVP framework of relational scalar-vector-potential fields provides a natural language for admissibility dynamics. In

RSVP, the state of a system is described by a triple  $(\Phi, \mathbf{v}, S)$  where:

- $\Phi$  is a scalar concentration field measuring the density of representational attention,
- $\mathbf{v}$  is a velocity field governing the flow of representational resources,
- $S$  is an entropy field measuring representational expenditure.

A trajectory is admissible within the RSVP framework when the flow generated by  $\mathbf{v}$  can stably transport  $\Phi$  through the region in question without uncontrolled growth of  $S$ . A persistent anomaly in RSVP terms appears as a region where  $S$  grows without bound under iterated application of the admissible repair dynamics. The scalar field  $\Phi$  concentrates on the anomaly—more and more representational attention is devoted to it—but the velocity field cannot generate a flow that dissolves the anomaly. This is the RSVP signature of an ontology boundary: the model is applying maximum representational resources and still failing.

The connection to the Persistence–Obstruction Schema is direct. The growth of  $S$  without stabilization corresponds to the non-vanishing of the obstruction class  $[\tau]$ . The admissibility boundary is the locus in state space where  $[\tau] \neq 0$ .

## CHAPTER 10

# Ontology Revision

### 10.1 When the Coordinate Chart is Exhausted

**Definition 10.1** (Ontology Revision). *Ontology revision* occurs when a learner reaches the ontology boundary of its current model  $\mathcal{M}$  and acquires a new model  $\mathcal{M}'$  such that:

- (i)  $\mathcal{A}(\mathcal{M}) \subsetneq \mathcal{A}(\mathcal{M}')$ ,
- (ii)  $x_{\text{obs}} \in \mathcal{A}(\mathcal{M}')$ ,
- (iii) the repair algebra  $\mathcal{R}'$  of  $\mathcal{M}'$  contains operations not in  $\mathcal{R}$ .

Ontology revision is not parameter updating. It is admissibility expansion: the new model can coherently represent trajectories that the old model could not.

### 10.2 Representational Inertia and the Rarity of Revision

Earlier framings of the framework might implicitly suggest that promotion is automatic: if an anomaly persists, it even-

tually triggers ontology revision. It does not. Promotion is expensive.

Replacing a typo costs almost nothing. Replacing a paragraph costs something. Replacing a theory may consume decades. Replacing an ontology may consume generations. The hierarchy has a built-in conservatism that is not a flaw. Without it, every anomaly would trigger catastrophic restructuring. The challenge is maintaining the correct balance between containment and promotion.

**Definition 10.2** (Representational Inertia). The *representational inertia*  $I_{\mathcal{M}}$  of a model is the accumulated investment in its current coordinate system: established results, trained practitioners, dependent theories, institutional infrastructure.

---

### Derivation 3: Promotion Through the Repair Hierarchy

Let  $C_k$  be the containment capacity of level  $k$ ,  $D_k$  the commitment cost of promoting to level  $k + 1$ ,  $B_{k+1}$  the expected benefit of persistence reduction at level  $k + 1$ , and  $I_{\mathcal{M}}$  the representational inertia. Promotion occurs when expected utility  $U_{k+1} > U_k$ , giving a promotion rule:

$$\Pr(\varepsilon : k \rightarrow k + 1) = \sigma(\alpha P_k(\varepsilon) - \beta C_k - \gamma D_k - \mu I_{\mathcal{M}}),$$

where  $\sigma$  is a logistic function. Promotion is likely when persistence is high, containment capacity is low, commitment cost is tolerable, and inertia is small. When  $I_{\mathcal{M}}$  is large, anomalies may be recognized and even promoted internally

while ontology revision remains blocked at the institutional or paradigmatic level.

---

The promotion equation explains several historical patterns. Mercury’s anomaly was recognized by Le Verrier in 1859. The ontology revision occurred in 1915. Fifty-six years elapsed because  $I_{\mathcal{M}}$  for Newtonian mechanics was enormous: every calculation in solar system astronomy depended on it, every observatory was equipped to apply it, every physicist had been trained in it. The anomaly was persistent and recognized. The inertia was too large for immediate revision.

### 10.3 The Dissolving Tear

A remarkable feature of ontology revision is that the persistent anomaly often vanishes entirely in the new framework—not by being “explained away” but by becoming a non-surprising trajectory.

**Definition 10.3** (Tear Dissolution). A non-abelian tear  $[\tau] \in H^1(\mathcal{M}; \mathcal{S}_{\mathcal{R}})$  is *dissolved* by an ontology revision to  $\mathcal{M}'$  if  $[\tau]$  maps to zero under the inclusion-induced map  $H^1(\mathcal{M}; \mathcal{S}_{\mathcal{R}}) \rightarrow H^1(\mathcal{M}'; \mathcal{S}_{\mathcal{R}'})$ .

Tear dissolution is what happened with Mercury. General relativity did not predict Mercury’s precession after the fact by adding a correction term. In pseudo-Riemannian space-time, the geodesic equation naturally produces the observed precession rate. The trajectory that was anomalous in Newtonian coordinates is the *expected* trajectory in Riemannian coordinates. The tear dissolved because the new coordinate system had room for it.

## 10.4 Ontological Saliency and Representational Divergence

**Definition 10.4** (Ontological Saliency). The *ontological saliency* of an observation  $x$  is

$$\text{Sal}_{\text{onto}}(x) = d_{\text{top}}(\mathcal{A}(\mathcal{M}_{\text{world}}), \mathcal{A}(\mathcal{M}_{\text{obs}}))$$

where  $\mathcal{A}(\mathcal{M}_{\text{world}})$  is the admissibility manifold of the actual world and  $\mathcal{A}(\mathcal{M}_{\text{obs}})$  is the admissibility manifold of the observer's model.

High ontological saliency means the observer's admissibility manifold and the world's admissibility manifold are diverging. Saliency, properly understood, is not surprise. It is evidence of impending coordinate failure.

---

### Derivation: Representational Divergence

Define the representational divergence at time  $t$  as

$$D_t = d(\mathcal{A}(\mathcal{M}_{\text{obs}}), \mathcal{A}(\mathcal{M}_{\text{world}})).$$

Then  $dD_t/dt$  measures the rate of ontology drift—the speed at which the observer's admissibility manifold is diverging from the world's. The persistence of an anomaly accumulates as

$$P(\varepsilon, \mathcal{R}) \approx \int_0^T \frac{dD_t}{dt} dt,$$

giving the following interpretation: persistent anomalies are accumulated evidence that the observer's admissibility manifold and the world's are diverging. The repair sequence is an

experiment that measures this divergence. The Admissibility Log is the instrument. The obstruction class is the result.

---

Part V

Learning and  
Representation

## CHAPTER 11

# Against Error Minimization

### 11.1 The Standard Account and Its Concealed Limits

The standard account of machine learning holds that a model improves by minimizing a loss function. Gradient descent, Bayesian updating, and reinforcement learning all share this structure. They differ in how they represent uncertainty and how they update parameters, but they agree that the goal is to reduce the discrepancy between predictions and observations as measured by some fixed criterion.

This account is not wrong. For a vast range of practical problems, error minimization produces excellent models [15, 39]. The critique developed here is not that error minimization fails, but that it succeeds in a way that systematically conceals its own limits. Modern optimization procedures are remarkably effective at reducing residuals while remaining largely agnostic to representational adequacy [45, 40].

## 11.2 What Error Minimization Cannot See

A loss function  $L : \Sigma \times \Sigma \rightarrow \mathbb{R}_{\geq 0}$  assigns a scalar to every prediction-observation pair. The optimization process moves through parameter space and asks at each step: which direction reduces the scalar? This formulation cannot, in principle, ask whether the scalar is the right thing to minimize. It cannot ask whether the discrepancy being measured is at Level 0 or Level 2 in the discrepancy hierarchy. It cannot track whether a residual has survived a sequence of repair operations or whether it appeared for the first time this epoch. It has no notion of the repair history of an anomaly.

Formally: a standard learning system equipped only with a loss function and a gradient oracle cannot compute  $\text{Pers}_{\mathcal{R}}(\varepsilon)$  for any non-trivial  $\mathcal{R}$ . It can compute instantaneous loss, but not the history of survival under repair. Related concerns appear in causal representation learning and model robustness research [32, 43].

## 11.3 Gradient Descent as Level-0 Repair

Every step of gradient descent is a Level-0 repair: a small numerical adjustment to parameters to reduce instantaneous loss. A single gradient step is not even a full repair operation in the sense defined in Part I; it is a fraction of one.

The consequence is that gradient-based learning cannot promote anomalies up the repair hierarchy. It applies Level-0 repairs indefinitely. An anomaly that is persistent at Level 0 but would reveal itself as a Level-2 topological discrepancy under proper repair triage is instead averaged over, diluted, or overfitted away. The topology remains invisible.

---

This is one explanation for the persistent difficulty of distribution shift in deep learning. A model trained on a distribution  $P$  develops a rich parameterization of  $P$ . When deployed on a shifted distribution  $Q$ , the anomalies often have topological content:  $M_Q \not\cong M_P$  as manifolds. But the training process has no mechanism for detecting this. It applies Level-0 repairs to what are Level-2 discrepancies. The model sees small average loss and interprets the situation as normal.

## CHAPTER 12

# Persistence-Driven Learning

### 12.1 Tracking the Repair History

A persistence-driven learner differs from a standard learner in one fundamental respect: it tracks the history of its repair attempts and uses survival records to assign priorities. The proposal is broadly compatible with information-theoretic and free-energy perspectives on adaptive systems [11, 1].

**Definition 12.1** (Persistence Monitor). A *persistence monitor* is an auxiliary system that:

- (i) records the repair level  $k_t$  applied at time  $t$ ,
- (ii) records the residual  $\varepsilon_t$  after each repair,
- (iii) computes an empirical persistence estimate by tracking survival across diverse repair classes,
- (iv) flags anomalies whose estimated persistence exceeds a threshold  $\theta$ .

The persistence monitor adds no parameters to the model and changes no gradients. It operates as a diagnostic layer on

top of the standard learning process, providing information that the standard process cannot see.

## 12.2 Persistence-Weighted Attention

In a learning system with attention mechanisms, the persistence monitor can directly influence the allocation of computational resources.

**Definition 12.2** (Persistence-Weighted Attention). Attention is *persistence-weighted* if

$$A(x) \propto f(\hat{P}(x))$$

where  $f$  is a monotone increasing function. The most important observations are those with highest estimated persistence, not highest instantaneous error.

Persistence-weighted attention reallocates resources toward observations where repair has repeatedly failed—toward the places where the model’s coordinate chart is most likely to be failing. It is a systematic implementation of repair triage.

## 12.3 Detecting Representational Boundaries During Training

A practical implication of persistence-driven learning is the possibility of detecting representational boundaries during training, before deployment-time failures reveal them.

During training, a persistence monitor tracks the survival history of anomalies across epochs. Anomalies that persist across many training epochs despite parameter updates are

candidates for Level-2 or higher discrepancies. A spike in the estimated persistence distribution—sudden appearance of many highly persistent small anomalies—indicates that the model is approaching the boundary of its representational capacity.

This is analogous to the physical signal preceding a topological phase transition: long-range correlations appear before the order parameter changes. Monitoring  $\hat{P}(\varepsilon)$  during training, in addition to standard loss  $L(\theta)$ , provides early warning of representational boundaries that loss monitoring cannot detect.

## CHAPTER 13

# Ontology-Expanding Learning

### 13.1 Learning as Representational Growth

**Definition 13.1** (Ontology-Expanding Learner). An *ontology-expanding learner* is a learning system equipped with:

- (i) a base model  $\mathcal{M}$  with admissible repair algebra  $\mathcal{R}$ ,
- (ii) a persistence monitor tracking  $\hat{P}(\varepsilon)$  for all active anomalies,
- (iii) an expansion trigger: a condition under which the learner initiates search for  $\mathcal{M}' \supset \mathcal{M}$ ,
- (iv) an expansion mechanism: a procedure for constructing extensions of the model architecture.

The expansion trigger fires when  $\hat{P}(\varepsilon) > \theta$  for some anomaly  $\varepsilon$  and some threshold  $\theta$ , indicating that the anomaly has survived enough repairs to be a candidate for structural diagnosis.

### 13.2 Persistent Anomalies as Generators

In the algebraic sense, persistent anomalies are generators of representational growth. Each persistent anomaly constrains the form that any adequate extension must take. The new model  $\mathcal{M}'$  must:

- (i) contain the admissibility manifold of  $\mathcal{M}$ ,
- (ii) include the observation  $x_{\text{obs}}$  in  $\mathcal{A}(\mathcal{M}')$ ,
- (iii) dissolve the obstruction class  $[\tau(\varepsilon)]$ .

These constraints narrow the search for  $\mathcal{M}'$  considerably. The anomaly is not just a problem to be solved; it is a specification of what the solution must achieve.

### 13.3 The Geometry of Repair Space

The framework developed in this monograph has been progressively clarifying what its ultimate object of study is. We began with anomalies. We moved to repair sequences. We arrived at persistence. But the deepest object is neither the anomaly nor the repair: it is the *geometry of repair space itself*.

**Definition 13.2** (Repair Landscape). The *repair landscape* of a model  $\mathcal{M}$  is the pair  $(\mathcal{M}, \mathcal{R})$  viewed as a geometric object: the model manifold  $\mathcal{M}$  equipped with the action of the repair algebra  $\mathcal{R}$  as a transformation group.

Two models may produce identical residuals and identical persistence invariants yet have dramatically different repair landscapes. One may have many short repair paths to

the correct ontology; another may require crossing enormous barriers. The repair landscape determines not just whether ontology revision is possible but how difficult it is to reach.

This reframes the entire enterprise. The monograph begins looking less like a theory of anomalies and more like a theory of navigation through representational spaces. Admissibility manifolds, reachability, CLIO projections, RSVP dynamics, and trajectory memory are all fundamentally concerned with the same question: given a current representation, what paths of transformation remain available? The anomaly is simply the thing that reveals where those paths begin to run out.

### 13.4 CLIO Projection Failures as Ontology Revision

In the CLIO framework of constraint-limited information ontology, projection failures occur when the operator  $\mathcal{P}$  mapping from the full state space to the observed subspace loses essential structure. A CLIO projection failure is precisely an ontology boundary: the projective compression has discarded information that is needed to continue the trajectory.

When a CLIO projection produces persistent anomalies in the projected description, those anomalies are evidence that the projection has discarded structure relevant to the dynamics. The ontology-expanding response is to enlarge the projection to restore the missing degrees of freedom. In RSVP terms: a CLIO projection failure is a region where the entropy  $S$  of the projected description grows faster than the velocity field  $\mathbf{v}$  can compensate. The projected description is becoming increasingly inadequate, and the failure is detectable as

a non-abelian tear between the full and projected trajectory manifolds.

## Part VI

# Suppressed Tears

## CHAPTER 14

# How Tears Get Hidden

### 14.1 The Dual of the Persistence Principle

The Persistence Principle tells us when structure is missing. But a dual and equally important phenomenon exists: structural anomalies that should be persistent are sometimes absorbed by repair operations that should not be able to absorb them. The tear is present but hidden.

**Definition 14.1** (Tear Suppression). A tear  $[\tau] \neq 0$  is *suppressed* if the repair sequence drives  $\|\varepsilon_n\| \rightarrow 0$  without dissolving the obstruction class. The model appears to fit the observations while containing an unresolved structural discrepancy.

Suppressed tears are the most dangerous failure mode for learning systems. The scalar loss looks good. The topology is wrong.

## 14.2 Three Mechanisms of Suppression

### 14.2.1 *Expressive Algebra Containment*

When a repair algebra at level  $k$  is sufficiently expressive, it can contain anomalies that properly belong at level  $k + 1$ . The repair reduces the metric manifestation of the anomaly to zero while leaving its topological invariant untouched.

This is the epicycle mechanism. Each epicycle is a perfectly valid Level-1 geometric repair. The problem is that the underlying discrepancy was categorical: the heliocentric structure of the solar system requires a different category of model, not a better fit within the geocentric category. But the epicycle system is expressive enough that the metric discrepancy can always be reduced to within observational tolerance. The tear is contained at Level 1 and never promoted.

### 14.2.2 *Commitment Asymmetry*

Higher-level repairs are more costly and partially irreversible. A rational agent with finite resources will require stronger evidence before attempting a higher-level repair. This creates a systematic bias: the evidence threshold for ontology revision is very high, meaning that even anomalies that have survived many lower-level repairs may not be promoted if the evidence for promotion is below the commitment threshold. Each individual repair attempt is rational; the aggregate sequence is collectively irrational because it prevents the accumulation of evidence that would justify a costlier intervention.

### 14.2.3 Promotion Blindness

As established in Part I, promotion requires memory. A system that cannot track the repair history of an anomaly cannot recognize that it has survived multiple repair levels. Each repair attempt is treated as a fresh encounter with the anomaly. The anomaly never climbs the hierarchy because the hierarchy is never traversed. This is the most fundamental form of suppression, because it operates not by neutralizing the evidence for structural problems but by preventing the accumulation of that evidence in the first place.

---

#### Derivation 4: Repair Containment and Overfitting

A repair level  $k$  *contains* an anomaly if there exists some  $R \in \mathcal{R}_k$  such that  $\|R(\varepsilon)\| \approx 0$  even though the structural obstruction remains:  $[\tau(\varepsilon)] \neq 0$ . This is the mathematical form of overfitting. The residual disappears; the tear remains. The model has repaired the measurement while preserving the representational failure. Containment probability is governed by the promotion equation: when  $C_k$  is large,  $\Pr(\varepsilon : k \rightarrow k + 1)$  collapses, and anomalies accumulate at Level 0 regardless of their structural depth.

---

## CHAPTER 15

# Epicycles and the Convergence Trap

### 15.1 The Ptolemaic System as a Case Study

The Ptolemaic system of epicycles is often presented as an example of scientific stubbornness: a wrong theory defended past its useful life. But from the perspective of the repair framework, it is better understood as a systematic and locally rational process of tear suppression. The historical interpretation follows the distinction between anomaly absorption and paradigm change emphasized by Kuhn and Lakatos [24, 25].

The core problem was that the model's representational architecture was genuinely inadequate: a geocentric coordinate system cannot represent the actual geometry of solar system dynamics without massive complexity. But each individual repair—each new epicycle, each adjusted equant—was locally effective. It reduced the metric discrepancy in the orbital predictions.

**Definition 15.1** (Repair Cascade). A *repair cascade* is a sequence of repair operations in which each repair generates a new anomaly that requires a further repair, producing an

expanding tree of local corrections rather than a converging sequence of residuals.

The Ptolemaic system underwent centuries of repair cascades. Each new observation triggered a new epicycle, which improved local fit but added complexity that generated new deviations at the boundaries of its validity. The repair sequence was not converging; it was branching. This is a diagnostic signature of a system containing a suppressed Level-2 or Level-3 tear. The repair entropy was exploding while the underlying obstruction class remained intact.

## 15.2 The Convergence Trap

**Definition 15.2** (Convergence Trap). A learning system is in a *convergence trap* when its repair sequence drives the observable residual toward zero while a non-trivial obstruction class persists in the model’s structure.

A diagnostic criterion for the convergence trap: if the repair sequence requires steadily increasing complexity to maintain a converging residual—if the size of the repair operations must grow while the residual shrinks—the system is likely in a convergence trap. Ptolemaic astronomy required more epicycles each decade. Modern neural networks require more parameters each year for comparable performance gains on challenging benchmarks. Both patterns are consistent with convergence-trap dynamics. The suppression of structural information by excessive model flexibility resembles familiar concerns in statistical learning theory [42, 15].

## CHAPTER 16

# Overfitting and the Ptolemaic Principle

### 16.1 Classical Overfitting and Its Topological Reinterpretation

Classical statistical learning theory presents overfitting as a variance problem: a model with too many parameters fits training data noise, producing poor generalization. The remedy is regularization. This account is correct as far as it goes, but it treats overfitting as a numerical phenomenon. We propose a complementary account: overfitting is a form of tear suppression.

When a neural network has more parameters than the effective dimensionality of the data manifold, the excess parameters constitute additional degrees of freedom in the Level-0 repair algebra. The network can use these degrees of freedom to absorb any anomaly at Level 0—any pointwise discrepancy—by adjusting parameters to locally fit the observation. Topological and categorical discrepancies that should be promoted are instead absorbed at Level 0.

**Conjecture 16.1** (Overfitting as Suppression). *Let  $\mathcal{M}_\theta$  be a*

*parametric model family with parameter space  $\Theta$ . If the dimension of  $\Theta$  exceeds the topological complexity of the data manifold, then for any observation with a non-trivial tear  $[\tau] \neq 0$ , there exists a sequence of parameter updates driving  $\|\varepsilon_n\| \rightarrow 0$  without dissolving  $[\tau]$ .*

Overfit models may therefore be topologically “wrong” in ways that their training performance entirely conceals. The model has learned to suppress the tear rather than to represent the structure it signals.

## 16.2 Generalization as Topological Compatibility

This analysis suggests a reinterpretation of generalization. A model generalizes well not merely when it is statistically consistent with the data distribution. It generalizes well when its trajectory manifold  $M_{\text{pred}}$  is *homotopy equivalent* to the world’s trajectory manifold  $M_{\text{world}}$ . Generalization failure occurs when  $M_{\text{pred}} \not\approx M_{\text{world}}$ —when the topological structure of the model’s predictions diverges from the topology of the world’s dynamics.

*Remark 16.2.* This reinterpretation connects the present framework to the literature on topological data analysis and geometric deep learning, where the homotopy type of data manifolds is increasingly recognized as an important quantity for generalization. The present framework provides a theoretical basis for why topology matters: not as a mathematical curiosity, but as the level at which suppressed tears live.

### 16.3 The Ptolemaic Principle

**Conjecture 16.3** (Ptolemaic Principle). *Let  $\mathcal{M}$  be a model with repair algebra  $\mathcal{R}$  of dimension  $d$ . If  $d$  exceeds the topological complexity of the true trajectory manifold, then for any persistent anomaly  $\varepsilon$  with  $\text{Pers}_{\mathcal{R}}(\varepsilon) > 0$ , there exists an overfit extension  $\mathcal{M}' \supset \mathcal{M}$  with repair algebra  $\mathcal{R}'$  such that  $\text{Pers}_{\mathcal{R}'}(\varepsilon) = 0$  but  $[\tau(\varepsilon)] \neq 0$  in  $H^1(\mathcal{M}'; \mathcal{S}_{\mathcal{R}'})$ .*

**Corollary 16.4.** *A system that monitors only metric residuals cannot distinguish genuine convergence (where  $[\tau] = 0$ ) from convergence-trap suppression (where  $[\tau] \neq 0$  but  $\|\varepsilon\| \rightarrow 0$ ). A system additionally monitoring  $[\tau]$ —via holonomy estimates or topological data analysis—can.*

The Ptolemaic Principle is the dual of the Persistence Principle. Persistence tells us when structure is missing. Suppression tells us why structure remains undiscovered. A sufficiently expressive repair algebra can always erase the metric evidence for a structural tear while leaving the tear itself intact.

The implications are immediate. First, convergence of a repair sequence is *not* evidence of ontological adequacy. Second, the history of epicycle-style science describes a failure mode in principle available to any learning system with a sufficiently flexible repair algebra—including modern deep learning systems. Third, the appropriate response to convergence is not celebration but diagnostic scrutiny: does the repair sequence converge by dissolving structural obstructions, or by absorbing them into increasingly complex local adjustments?

Part VII

Applications

## CHAPTER 17

# Mercury and General Relativity

### 17.1 The Anomaly

Mercury's perihelion precesses by 5600 arcseconds per century. Perturbations from other planets account for approximately 5557 arcseconds. The residual of 43 arcseconds per century was documented by Le Verrier in 1859 [44]. For the next fifty-six years, this anomaly climbed every level of the Newtonian repair hierarchy.

#### *17.1.1 Level-0 Repairs*

Measurement improvements refined the precession value but did not change the residual's sign or order of magnitude. The anomaly was not an observational artifact. If anything, its precision increased as instrumentation improved—a signature of amplified modes in the persistence spectrum.

#### *17.1.2 Level-1 Repairs*

Parameter adjustments within Newtonian gravity—changing the mass of the Sun, adjusting the oblateness, modifying plan-

etary perturbations—were studied extensively and found insufficient. Hall proposed in 1894 that modifying the inverse-square law to a slightly different power could account for the precession, but this modification broke agreement with other planetary observations.

### *17.1.3 Level-2 Repairs*

Hypothetical new bodies were proposed: a planet interior to Mercury's orbit (Vulcan), distributed interplanetary matter, an unseen population of small bodies. Extensive searches found nothing. Each failed repair increased the admissibility pressure without reducing the anomaly's invariant modes.

### *17.1.4 The Ontology Revision*

Einstein's field equations, applied to the two-body problem, yield a geodesic equation with an additional term producing a perihelion precession of [8]

$$\Delta\phi = \frac{6\pi GM}{c^2 a(1 - e^2)}$$

per orbit. For Mercury, this gives precisely 43 arcseconds per century. The calculation requires no free parameters; it is a direct consequence of the background geometry.

## **17.2 What the Persistence Framework Adds**

The standard narrative presents this as a story about a wrong theory eventually replaced by a right one. The persistence framework adds precision to two aspects.

First, the anomaly was not evidence against Newtonian gravity in the scalar sense—the vast majority of Newtonian predictions remained accurate. The anomaly was evidence of a non-trivial obstruction class in the trajectory manifold: a region of state space where Newtonian coordinates were topologically inadequate.

Second, the resolution did not *fit* the anomaly. It *dissolved* it. In pseudo-Riemannian spacetime, Mercury's trajectory is a geodesic. There is no residual, no correction term, no patch. The tear was dissolved by expanding the admissibility manifold to include curved geometries. This is ontology revision in the precise sense: the obstruction class maps to zero under inclusion into the Riemannian framework.

## CHAPTER 18

# Scientific Revolutions as Ontology Revision

### 18.1 A Persistence-Based Account of Scientific Change

Kuhn's account of scientific revolutions distinguishes normal science—puzzle-solving within a paradigm—from revolutionary science—replacement of the paradigm itself [24]. Lakatos's research programmes provide a natural interpretation of repair hierarchies [25]. The role of failed predictions in theory revision reflects Popper's falsificationist tradition [35]. The persistence framework provides a mathematical account of what distinguishes the normal from the revolutionary.

Normal science is Level-0 through Level-2 repair: parameter adjustment, coordinate refinement, mechanism addition. Revolutionary science is Level-3 and above: categorical revision, ontology expansion. The persistence monitor tells us when the transition is necessary: when anomalies have survived the full normal-science repair hierarchy. At that point, further investment in normal-science repair is governed by the Ptolemaic Principle—it will either continue to fail or succeed by suppression.

## 18.2 The Role of Anomaly Persistence in Discovery

The historical record suggests that the most important scientific anomalies share the following properties:

- (i) They are numerically small relative to the total predictive success of the prevailing theory.
- (ii) They are documented over long time periods, establishing empirical persistence.
- (iii) They survive an unusually wide variety of repair attempts within the existing framework.
- (iv) Their persistence spectrum contains invariant or amplified modes: as the prevailing theory becomes more precise, the anomaly becomes sharper, not smaller.
- (v) Their eventual resolution requires a concept absent from the existing framework.

The ultraviolet catastrophe is the clearest case. The deviation between classical prediction and observation was only severe at high frequencies; at low frequencies, classical theory worked well. The anomaly persisted through decades of repairs until Planck introduced energy quantization [33]. The new concept (the quantum of action) had no analogue in classical physics. The tear required not a better classical theory but a different kind of theory.

## 18.3 Science as Collective Triage

From the persistence perspective, science is a collective repair triage process. The scientific community functions as a

distributed repair hierarchy, with different communities and traditions performing repairs at different levels. A new observation enters at Level 0; it is checked against measurement theory first, then against existing models, then against the space of possible models, then against the conceptual framework.

The community's collective persistence monitor is the literature: the record of failed repair attempts. An anomaly that has generated decades of failed repair papers is an anomaly with high documented persistence—a rich Admissibility Log maintained collectively. The community gradually recognizes that no available repair will work, and the conditions for paradigm change are established.

The failure mode is also visible at the collective level: a community that lacks memory of its repair history will repeat past repair attempts without recognizing them as past. The literature performs the memory function, but only when it is consulted. A community that fails to search existing literature before attempting repairs is operating with reduced memory and therefore reduced capacity for persistence detection—the collective form of promotion blindness.

## CHAPTER 19

# Biological Adaptation

### 19.1 Evolution as Persistent Anomaly Processing

Biological evolution can be viewed through the persistence framework as follows. An organism is a model of its environment: its physiology predicts what the environment will provide and adjusts behavior accordingly. Environments change. Some changes are Level-0 perturbations: temperature fluctuations, prey abundance fluctuations, that can be absorbed by physiological and behavioral flexibility. Others are structural: the environment's topology has changed in ways that no existing behavioral or physiological repair can address.

Genetic mutation and selection constitute a higher-level repair algebra. When phenotypic flexibility fails to absorb an environmental anomaly, the anomaly persists across generations. Selection pressure is a persistence signal: high mortality in response to environmental conditions that cannot be behaviorally avoided is the biological equivalent of a promoted anomaly. The repair hierarchy in biology extends from behavioral adjustment (Level 0) through physiological acclimation (Level 1) to genetic mutation (Level 2) and develop-

mental program revision (Level 3).

## 19.2 Speciation as Ontology Revision

Speciation—the evolution of a new species—is a biological ontology revision. The new species carries a representational architecture (phenotype, developmental program, niche specification) that extends the admissibility manifold of the lineage to include environmental conditions that the ancestor could not inhabit.

The persistence of anomalies—environmental pressures that killed individuals at high rates without behavioral remedy—is the generator of selection pressure that drives speciation. In the persistence framework, speciation is what happens when a biological lineage’s repair hierarchy is exhausted and the only remaining response is admissibility expansion. The new species does not merely tolerate the previously anomalous environment; it inhabits it as a normal part of its admissibility manifold.

## 19.3 Developmental Regulation

At the cellular level, developmental regulation involves ongoing repair triage. A developing organism faces continuous small discrepancies between target morphology and current state. Most are resolved by Level-0 biochemical adjustments. Occasionally a discrepancy persists across cell cycles and regulatory checkpoints, which may trigger a Level-3 response: a change in the developmental program itself.

Persistent developmental anomalies—homeotic transformations, regulatory failures—are often the most evolutionar-

ily productive mutations. They involve changes not to individual gene products but to the regulatory architecture that governs development. They are, in the present framework, categorical changes to the biological model. This corresponds to the bottom-right cell of the discrepancy matrix: ontology boundary anomalies that signal that the category of developmental programs currently available to the lineage is inadequate to the environmental challenge.

## CHAPTER 20

# Machine Learning and Representation

### 20.1 Foundation Models and the Persistence Problem

Foundation models trained on large corpora demonstrate strong average performance across many tasks. Their training is a massive repair operation: gradient descent applied to the full dataset drives the training loss toward zero. The model's parameter space is large enough that the Ptolemaic Principle applies in its conjectured form: essentially any training-distribution anomaly can be absorbed at Level 0.

The consequence is that foundation models are optimally prone to convergence-trap suppression. Their expressive repair algebras absorb evidence of structural discrepancies. Persistent anomalies—systematic failures that survive many training iterations—are present but invisible to standard monitoring.

The out-of-distribution failure problem is a manifestation of this. When a foundation model is deployed on inputs outside its training distribution, the metric distance is small for many inputs, but the topological structure of the out-

of-distribution manifold may be substantially different from the training manifold. The model’s tear has been suppressed during training; it appears during deployment when the suppression mechanism (training distribution repair) is no longer available.

## 20.2 Representation Learning and Topological Fidelity

The goal of representation learning is to construct a latent space that preserves the structure of the data. From the persistence perspective, “preserving structure” means achieving homotopy equivalence between the data manifold and the representation manifold:

$$M_{\text{data}} \simeq M_{\text{repr}}.$$

A representation that achieves this has dissolved all tears. A representation that merely minimizes reconstruction error may have suppressed tears rather than dissolved them.

This motivates a topological approach to representation learning: instead of monitoring reconstruction loss, monitor the homology of the representation manifold and compare it to the estimated homology of the data manifold. Discrepancies in homology indicate persistent tears. Topological data analysis provides tools for this; the persistence framework provides the theoretical justification for why it should be done.

## 20.3 A Research Agenda

The persistence framework suggests several concrete directions for machine learning research.

**Persistence monitoring during training.** Track the repair history of training anomalies across epochs. Anomalies that persist across many epochs are candidates for structural diagnosis. A persistence spike during training—sudden appearance of many highly persistent small anomalies—is an early warning of impending representation failure.

**Commutator-based diagnostics.** Systematically apply pairs of repair operations in both orders. Persistent  $[R_a, R_b](\varepsilon) \neq 0$  across diverse repair families indicates structural tears rather than slow convergence.

**Holonomy-based data augmentation.** Generate training examples by transporting existing examples around loops in data space. If the model's holonomy around these loops is non-trivial, the resulting examples probe the regions near non-abelian tears. Targeted augmentation in these regions may help dissolve rather than suppress structural discrepancies.

**Topology-regularized training.** Add a regularization term penalizing discrepancies between the homology of the model's prediction manifold and estimated data homology. This would directly discourage suppression in favor of dissolution.

**Repair-entropy monitoring.** Track the ratio  $E_R = -\Delta p_n / \Delta S_R$ . Declining repair efficiency is a leading indicator of convergence-trap dynamics.

**Ontology-expanding architectures.** Design architectures capable of adding new representational dimensions when persistence monitors indicate boundary proximity. Neural architecture search driven by persistence signals rather than validation loss.

## CHAPTER 21

# RSVP, CLIO, and Admissibility Fields

### 21.1 The RSVP Framework

The Relational Scalar-Vector-Potential (RSVP) framework describes physical and representational dynamics in terms of a triple  $(\Phi, \mathbf{v}, S)$ , where  $\Phi$  is a scalar concentration field,  $\mathbf{v}$  a velocity field, and  $S$  an entropy field. The dynamics of this triple are governed by a system of coupled field equations that describe how representational attention flows, concentrates, and dissipates.

The RSVP framework provides natural language for admissibility dynamics. The admissibility manifold  $\mathcal{A}(\mathcal{M})$  corresponds to the region of field configuration space where the RSVP dynamics generate stable trajectories of  $\Phi$  under the flow of  $\mathbf{v}$ . A trajectory is admissible when the entropy field  $S$  remains bounded along it.

## 21.2 Non-Abelian Tears in RSVP

In RSVP dynamics, a non-abelian tear appears as a region where:

- (i) The scalar field  $\Phi$  concentrates: the model is devoting increasing attention to the region.
- (ii) The velocity field  $\mathbf{v}$  fails to generate a stable flow: the transport of  $\Phi$  is becoming incoherent.
- (iii) The entropy field  $S$  grows without bound: the representational cost of tracking the trajectory is escalating.

This triple signature—concentration without coherent flow and unbounded entropy growth—is the RSVP signature of an ontology boundary. The model is approaching the edge of its coordinate chart.

The holonomy discrepancy of the RSVP connection around a loop  $\ell$  measures the accumulated failure of the velocity field to maintain coherent transport. A region with high holonomy discrepancy is a region near a non-abelian tear, and corresponds directly to the non-vanishing of the obstruction class under the Persistence–Obstruction Schema.

## 21.3 CLIO Projection and Representational Compression

The CLIO framework governs the projection from full state space to observable subspaces. A CLIO projection is admissible when it preserves the relevant topological structure: when the projection map induces an equivalence on the homotopy classes of trajectories.

A CLIO projection failure occurs when this condition breaks down: when the projected description cannot represent the loops in trajectory space that carry non-trivial holonomy. The projection has discarded the dimensions needed to represent the tear.

In the persistence framework, a CLIO projection failure is a specific type of ontology revision problem: the model has projected itself into a subspace that is too small to contain the full repair algebra needed to address the anomaly. The response is to expand the projection—to include more dimensions in the CLIO output—until the relevant topological structure is preserved.

## 21.4 Lamphrodyne Relaxation and Tear Dissolution

The lamphrodyne relaxation process in RSVP describes the gradual equilibration of the entropy field following a perturbation. During relaxation, the velocity field reorganizes to restore coherent transport of the scalar concentration field.

In the persistence framework, lamphrodyne relaxation is a Level-1 geometric repair: adjusting the flow structure to improve trajectory tracking. Successful relaxation corresponds to the anomaly being absorbed at Level 1. Failed relaxation—where the entropy field grows or oscillates rather than equilibrating—is the RSVP signature of a persistent anomaly that requires higher-level repair.

The RSVP dynamics thus provide an implementable instantiation of the abstract repair hierarchy. The level of the repair hierarchy reached by an RSVP anomaly is indicated by the behavior of the relaxation process: whether it con-

verges (Levels 0–1), oscillates (Level 2), or diverges (Level 3 and above).

## Conclusion: What Anomalies Tell Us

It is not the magnitude of the discrepancy but  
its stubbornness that matters.

---

*—the argument of this book*

We began with Mercury. We end with a general principle about the geometry of learning.

The argument of this monograph has proceeded through several stages that are worth recapitulating in reverse order, because the reversal is illuminating.

The deepest object of study is the geometry of repair space: the repair landscape  $(\mathcal{M}, \mathcal{R})$  that describes not just what a model predicts but what transformations are available to it and how difficult they are to perform. Most theories of learning are theories of the model. This framework is a theory of the model-and-its-transformations.

The key quantity in the repair landscape is persistence: the survival of an anomaly under all available repairs, measured relative to the algebra  $\mathcal{R}$  available to the observer. Persistence is not an intrinsic property of the anomaly. It is a relational property of the pair  $(\varepsilon, \mathcal{R})$ . The same phenomenon looks transient to a sufficiently expressive system and persistent to a simpler one. The same anomaly that climbs a

Newtonian repair hierarchy is absorbed instantly by general relativistic machinery once that machinery is available.

The mechanism by which persistence is computed is memory. Without a remembered repair trajectory, no anomaly can be recognized as persistent. The dependency chain runs backward from what is usually expected: memory enables persistence detection; persistence detection enables promotion; promotion enables ontology revision; ontology revision is the deepest form of learning. Memory is not a storage system for the outputs of learning. It is the computational infrastructure that makes the highest levels of learning possible at all.

The dual failure mode—the Ptolemaic Principle—is equally important. A sufficiently expressive repair algebra can always suppress the metric evidence for a tear without dissolving it. Overfitting is tear suppression. Bureaucratic drift is tear suppression. Paradigm defense under experimental pressure is tear suppression. In each case, the anomaly has not disappeared. It has been contained at a level that cannot address it, consuming resources while preventing promotion.

The history of science can be read as a long series of anomalies being sorted by their persistence. Most anomalies are transient, resolved at Level 0 or Level 1, and promptly forgotten. They were, in retrospect, uninteresting. A small number are persistent. They climb the repair hierarchy until no admissible repair remains. At that point they function not merely as unresolved residuals but as specifications: they tell us, by the structure of their persistence spectrum and their obstruction class, what the next theory must contain.

*An abelian error can be corrected.*

*A non-abelian tear requires the model to grow.*

Saliency, properly understood, is not the magnitude of surprise. It is evidence that the current representation lacks the degrees of freedom necessary to explain what happened. The most valuable anomalies in science have rarely been the largest ones. They have been the smallest ones that refused to disappear—the ones whose persistence spectrum contained invariant or amplified modes, whose commutators remained non-zero under every available pair of repair operations, whose admissibility pressure climbed through decades of failed attempts.

These are the places where the world was informing its models that they needed to become larger. The anomaly is the instrument by which a model discovers the shape of its own boundary. The repair sequence is the experiment. The Admissibility Log is the record. The geometry of repair space is the object. And the anomaly was being patient.

\* \* \*

**Persistence Principle** (Final Statement). Let  $\mathcal{M}$  be a model and  $\mathcal{R}$  its admissible repair algebra. An anomaly  $\varepsilon$  satisfying  $\text{Pers}_{\mathcal{R}}(\varepsilon) > 0$  is not a fact about the world. It is a fact about the geometry of the model used to represent the world. Its persistence invariant measures the distance between the model's admissibility manifold and the world's, integrated over the repair trajectory.

Persistent anomaly  $\implies$  missing structure.

The repair sequence is the experiment. The Admissibility Log is the record. The geometry of repair space is the object. And the anomaly is the instrument by which the model discovers the shape of its own boundary.

# Bibliography

## Bibliography

- [1] Amari, S. *Information Geometry and Its Applications*. Springer, 2016.
- [2] Awodey, S. *Category Theory*. Oxford University Press, 2010.
- [3] Bishop, C. M. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [4] Bott, R. and Tu, L. *Differential Forms in Algebraic Topology*. Springer, 1982.
- [5] Cartwright, N. *How the Laws of Physics Lie*. Oxford University Press, 1983.
- [6] Cover, T. M. and Thomas, J. A. *Elements of Information Theory*. Wiley, 2006.
- [7] De Lellis, C. *Rectifiable Sets, Densities and Tangent Measures*. European Mathematical Society, 2017.
- [8] Einstein, A. The foundation of the general theory of relativity. *Annalen der Physik*, 49:769–822, 1916.
- [9] Engelking, R. *General Topology*. Heldermann Verlag, 1989.

- 
- [10] Feyerabend, P. *Against Method*. Verso, 1975.
- [11] Friston, K. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.
- [12] Gallier, J. and Quaintance, J. *Differential Geometry and Lie Groups*. Springer, 2020.
- [13] Giere, R. *Explaining Science*. University of Chicago Press, 1988.
- [14] Goldblatt, R. *Topoi: The Categorical Analysis of Logic*. Dover, 2006.
- [15] Goodfellow, I., Bengio, Y., and Courville, A. *Deep Learning*. MIT Press, 2016.
- [16] Gromov, M. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhäuser, 1999.
- [17] Hanson, N. R. *Patterns of Discovery*. Cambridge University Press, 1958.
- [18] Hartshorne, R. *Algebraic Geometry*. Springer, 1977.
- [19] Hatcher, A. *Algebraic Topology*. Cambridge University Press, 2002.
- [20] Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- [21] Holland, J. H. *Adaptation in Natural and Artificial Systems*. MIT Press, 1992.

- 
- [22] Jaynes, E. T. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.
- [23] Kashiwara, M. and Schapira, P. *Categories and Sheaves*. Springer, 2006.
- [24] Kuhn, T. S. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962.
- [25] Lakatos, I. *The Methodology of Scientific Research Programmes*. Cambridge University Press, 1978.
- [26] Leinster, T. *Basic Category Theory*. Cambridge University Press, 2014.
- [27] Mac Lane, S. *Categories for the Working Mathematician*. Springer, 1971.
- [28] MacKay, D. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.
- [29] Michelson, A. A. and Morley, E. W. On the relative motion of the Earth and the luminiferous ether. *American Journal of Science*, 34:333–345, 1887.
- [30] Milnor, J. *Characteristic Classes*. Princeton University Press, 1974.
- [31] Nakahara, M. *Geometry, Topology and Physics*, 2nd ed. Taylor & Francis, 2003.
- [32] Pearl, J. *Causality*. Cambridge University Press, 2009.
- [33] Planck, M. On the law of distribution of energy in the normal spectrum. *Annalen der Physik*, 4:553–563, 1901.

- 
- [34] Polanyi, M. *The Tacit Dimension*. University of Chicago Press, 1966.
- [35] Popper, K. *The Logic of Scientific Discovery*. Routledge, 1959.
- [36] Quine, W. V. O. *Word and Object*. MIT Press, 1960.
- [37] Riehl, E. *Category Theory in Context*. Dover, 2016.
- [38] Rudin, C. Stop explaining black box machine learning models for high stakes decisions. *Nature Machine Intelligence*, 1(5):206–215, 2019.
- [39] Shalev-Shwartz, S. and Ben-David, S. *Understanding Machine Learning*. Cambridge University Press, 2014.
- [40] Simon, H. A. *The Sciences of the Artificial*. MIT Press, 1969.
- [41] Spanier, E. *Algebraic Topology*. McGraw-Hill, 1966.
- [42] Vapnik, V. *Statistical Learning Theory*. Wiley, 1998.
- [43] Verma, V. and Pearl, J. A theory of causal representation learning. *Proceedings of AAAI Workshops*, 2018.
- [44] Will, C. M. *Theory and Experiment in Gravitational Physics*. Cambridge University Press, 2014.
- [45] Wimsatt, W. *Re-Engineering Philosophy for Limited Beings*. Harvard University Press, 2007.