Hidden Curvature

Constraint Geometry, Projection,
and the Reconstruction of Meaning

Flyxion

Independent Researcher

2026

Abstract

This monograph develops a field-theoretic account of semantic meaning grounded in admissibility geometry rather than representational fidelity. The central claim is that meaning is not a property of representations but an invariant of the admissibility field—the structured space of constraint-governed transitions through which semantic trajectories move. Representations are projections of that field. What has been described in the literature as semantic collapse is not a failure of representational precision but the observable trace of a geometric phenomenon: the destruction of curvature by many-to-one projection. The intellectual center of the work is a reconstruction program: given the observable failures of representation, recover the hidden constraint geometry that produced them.

The argument proceeds in four stages. First, we construct the admissibility manifold X of semantic trajectories and equip it with the Fisher information metric induced by the conditional density $p(y \mid x)$ over admissible continuations. This metric is positive semi-definite everywhere and degenerates precisely at semantic bottlenecks, making it well-defined at the locations where existing Hessian-based approaches fail. Second, we introduce the vertical-horizontal decomposition of the projection $\pi : X \to M$ onto a representational manifold M and define the projected sectional curvature $\kappa_\pi(m)$ as the integrated mixed curvature over the fiber $\pi^{-1}(m)$. Third, we prove the Projection-Collapse Principle in both a weak and a strong form: observable representational mixing $\Lambda(m)$ is bounded below by a monotone function of $\kappa_\pi(m)$, and under regularity conditions the two quantities are asymptotically equivalent. This establishes that collapse measurements are geometric observables of the hidden admissibility geometry. Fourth, we formulate the inverse problem: given a field of mixing measurements over M, reconstruct the curvature distribution of the admissibility manifold from its projected trace.

Connections are developed to information geometry, retarded functional differential equations, and the field theory of synthetic cognition. The monograph concludes with a discussion of constraint-first semantic architectures, collapse-based auditing, and the philosophical consequences of treating constraint as ontologically prior to representation.

Keywords: admissibility geometry, semantic collapse, Fisher information metric, projection-collapse principle, inverse geometry, RSVP field theory, constraint-first semantics.

Contents

Abstract **i**

I The Phenomenon 1

1 What Collapse Looks Like **2**

1.1 Representational Systems and Their Failures 2
1.2 Operator Sensitivity as the First Diagnostic 2
1.3 The Phenomenology of Collapse Events 3
1.4 Collapse as a Signal 4

2 Existing Frameworks and Their Limits **5**

2.1 Embedding-Based Accounts 5
2.2 Modal Logic Approaches 5
2.3 The Regress Problem 6
2.4 The Inversion: Constraints Before Representations 7

II The Geometry 8

3 The Admissibility Manifold **9**

3.1 Semantic Trajectories as the Primary Objects 9
3.2 The Admissible Continuation Space 9
3.3 Local Charts from Admissibility Profiles 10
3.4 Regularity Conditions 11
3.5 Bottlenecks, Constraints, and Degeneracies 11

4 Tangent Spaces and Admissibility Flows **13**
4.1 Statistical Tangent Vectors . . . 13
4.2 Score Functions as Natural Coordinates . . . 14
4.3 The Fisher Inner Product . . . 14
4.4 Uniqueness of the Fisher Metric . . . 15
4.5 Semantic Interpretation of Admissibility Flows . . . 16

5 Admissibility Geodesics **17**
5.1 The Geodesic Energy Functional . . . 17
5.2 The Geodesic Equation . . . 18
5.3 Fisher Geodesic Distance . . . 18
5.4 Semantic Interpretation: Geodesics as Minimal Deformation . . . 19
5.5 Bottleneck Deflection and Escape Velocity . . . 19

6 Volume Growth and Accessibility Geometry **20**
6.1 Fisher Balls and Accessibility Volume . . . 20
6.2 Comparison Geometry: The Bishop-Gromov Theorem . . . 21
6.3 Curvature and Accessibility Volume . . . 21
6.4 Semantic Bottlenecks as Volume Suppression . . . 22

7 The Fisher Information Metric **23**
7.1 Why the Hessian Fails . . . 23
7.2 The Fisher Information Metric . . . 23
7.3 Derivation from Infinitesimal Admissibility Divergence . . . 24
7.4 Positive Semi-Definiteness and Its Semantic Meaning . . . 25
7.5 Connection to Information Geometry . . . 26
7.6 Curvature of the Admissibility Manifold . . . 26

8 Projections and the Vertical-Horizontal Decomposition **28**
8.1 Representations as Projections . . . 28
8.2 The Vertical-Horizontal Decomposition . . . 28

8.3 Projected Sectional Curvature 29
8.4 The O'Neill Tensors and Collapse-Driving Curvature 30
8.5 The Fiber as a Hidden Variable 31

III The Theorem 32

9 The Representational Mixing Functional 33
9.1 Beyond Operator-Specific Leakage 33
9.2 Definition of the Mixing Functional 33
9.3 Existing Diagnostics as Special Cases 34
9.4 Regularity Conditions on the Functional 35

10 The Projection-Collapse Principle 36
10.1 Fiber Divergence: The Bridge Between Curvature and Leakage 36
10.2 The Curvature-Divergence Inequality 37
10.3 Statement: Weak Form 38
10.4 Proof of the Weak Form 38
10.5 Statement: Strong Form 39
10.6 The Regularity Conditions and Their Semantic Content 40
10.7 Interpretive Consequences 40

IV The Inverse Problem 42

11 Collapse as Geometric Tomography 43
11.1 The Inverse Problem Formulated 43
11.2 Analogy with Geometric Tomography 43
11.3 Resolution Limits and the Cramér-Rao Bound 44
11.4 What Can and Cannot Be Recovered 44
11.5 The Reconstruction Operator 45

12 Spectral Reconstruction Theory 47

12.1 The Linearized Forward Operator 47
12.2 Singular Value Decomposition 48
12.3 Curvature Recovery via Truncated SVD 48
12.4 Tikhonov Regularization 49
12.5 The Singular Value Spectrum as Semantic Resolution Limit 50
12.6 Triplet Measurements as Curvature Probes 50
12.7 Entropy Measurements as Curvature Integrals 51
12.8 Perturbation Experiments as Tangent Vector Probes 52
12.9 Combining Probes: Toward Tomographic Reconstruction 52

13 Connection to Reentrant Value Field Theory 53

13.1 Bohlen's Framework and the Admissibility Manifold 53
13.2 The Small-Gain Condition as a Collapse Surface 53
13.3 The Lyapunov-Krasovskii Functional as Admissibility Density 54
13.4 Blueprint Admissibility as Constraint Geometry 55

14 Semantic Field Theory 56

14.1 From Diagnostics to Dynamics 56
14.2 RSVP as a Field-Theoretic Framework 56
14.3 Meaning as a Field-Theoretic Invariant 59

15 Variational RSVP Geometry 60

15.1 The Admissibility Action 60
15.2 Euler-Lagrange Equations: The Admissibility Field Equation 61
15.3 The Mexican-Hat Potential and Semantic Phase Transitions 61
15.4 Fisher-RSVP Correspondence: The Metric from the Action 62
15.5 Curvature Evolution Under RSVP Dynamics 63

16 Persistent Admissibility Topology 64

16.1 The Admissibility Filtration 64
16.2 Homology of the Admissibility Filtration 65

16.3 Persistence Diagrams 65
16.4 Topological Bottlenecks and Disconnected Basins 66
16.5 Semantic Topological Invariants 66

V Architecture and Infrastructure 68

17 Constraint-First Semantic Architectures 69
17.1 The Design Inversion 69
17.2 Admissibility Filtering as a Primitive Operation 70
17.3 Typed Representations and Projection Discipline 70

18 Diagnostics, Auditing, and Certification 72
18.1 The Diagnostic Hierarchy 72
18.2 Collapse Graphs as Audit Artifacts 72
18.3 Certification via Proof Obligations 73

VI Foundations and Consequences 74

19 Accessibility Fields, Trajectories, and Semantic Action 75
19.1 Accessibility Fields 75
19.2 Accessibility Entropy 76
19.3 Trajectory Primacy and Semantic Equivalence 76
19.4 Constraint Basins and Semantic Objects 77
19.5 Semantic Renormalization 78
19.6 The Semantic Action Functional 79

20 Intellectual Foundations and Related Frameworks 81
20.1 Information Geometry 81
20.2 Riemannian Geometry and Submersion Theory 82
20.3 Optimal Transport 82
20.4 Information Bottlenecks and Compression 83

20.5 Topological Data Analysis 83
20.6 Geometric Deep Learning 83
20.7 Constraint-Based Accounts of Cognition 84
20.8 Variational Field Theory 84
20.9 Process Philosophy and Structural Ontology 85
20.10 Positioning Relative to Existing Frameworks 85
20.11 The Ontological Status of Admissibility 86
20.12 Kinematics and Dynamics of Admissibility 87
20.13 Semantic Event Horizons 88
20.14 Discrete Admissibility Geometry 89
20.15 Meaning Without Representation 91
20.16 The Chinese Room Revisited 91
20.17 Indexicality and the Admissibility Manifold 92

21 Open Problems **93**
21.1 Foundational Problems 93
21.2 Geometric Problems 93
21.3 Dynamical Problems 94
21.4 Computational Problems 94

Conclusion **96**

A Information Geometry: Background and Notation **97**

B Differential Geometry: Curvature and Submersions **98**

C The O'Neill Formulas **99**

D Proof Details for the Weak Form **101**

E Connection to RSVP Field Theory **102**

Part I

The Phenomenon

Chapter 1

What Collapse Looks Like

1.1 Representational Systems and Their Failures

A representational system encodes structured information in a compressed space. The compression is useful: it allows a finite device to operate over an indefinitely large domain of possible states by treating structurally similar states as equivalent. But compression is also lossy. Some information is discarded, and the question that motivates this monograph is which information, why, and what the loss reveals about the hidden structure the representation was attempting to encode.

The failures that matter most are not random. When a representational system loses information, it does not do so uniformly across all distinctions. Certain classes of distinctions are systematically more vulnerable than others. These are the distinctions that depend not on the local features of a token or state but on its position within a structured space of possibilities: on what transitions from that position are admissible, what futures remain accessible, and what constraints govern the trajectory forward. A system that compresses by geometric proximity—treating nearby states as equivalent—will erase these distinctions whenever the relevant structure is not preserved by the proximity relation.

The phenomenon of semantic collapse names this systematic erasure. It is not a description of noise or approximation error. It is a description of a structural failure: the destruction, by compression, of the distinctions on which reasoning depends.

1.2 Operator Sensitivity as the First Diagnostic

The most immediate evidence of collapse comes from operator-sensitive constructions. Modal operators, epistemic operators, indexicals, and agentive constructions share a common feature: their meaning depends not on the intrinsic properties of their arguments but on the structural relations those arguments bear to a space of possibilities. The sentence “it must be raining” differs from “it is raining” not in its surface vocabulary but in the class of possible worlds over which it is evaluated. The sentence “she knows the door is locked” differs from “the door is locked” not in its propositional content but in whether that content is embedded within an epistemic frame.

Representational systems that treat semantic similarity as geometric proximity are systematically blind to these distinctions. They cluster tokens by co-occurrence and distributional context, which is sensitive to surface form but not to modal scope. The result is that necessity is blurred with truth, epistemic attitude is flattened into assertion, and indexical anchoring dissolves into description. The operators that structure logical inference are the first casualties of compression.

This is not a defect of any particular architecture. It is a consequence of the projection. Any compression that identifies states by their positional neighborhoods rather than by their accessibility structure will produce operator collapse.

1.3 The Phenomenology of Collapse Events

Collapse manifests across multiple domains in recognizably similar patterns. In linguistic systems, it appears as hallucination: the generation of fluent outputs that are semantically indistinguishable from correct ones in the compressed representation but logically incompatible with the constraints governing the original. In formal reasoning systems, it appears as scope errors: the confusion of necessity with contingency, of universal quantification with existential, of an agent’s action with an event that merely occurs. In perceptual systems, it appears as category boundary degradation: the failure to maintain separability between classes that are geometrically close in representation space but functionally distinct in the constraint structure.

The surface forms of these failures differ. Their underlying geometry is the same. In each case, a many-to-one projection maps a curved admissibility structure onto a flatter representational space, and the curvature destroyed by the projection reappears as confusions in the output. The specific content of the confusion depends on which curvature was destroyed. The structural form of the confusion is invariant.

1.4 Collapse as a Signal

The reorientation that drives the rest of this monograph is simple to state and far-reaching in its consequences. Collapse events are not errors to be corrected. They are measurement signals. A system that collapses in a particular pattern is revealing something about the hidden structure it is compressing. The pattern of collapse is a projected trace of the curvature of the admissibility manifold.

This is not a metaphor. The Projection-Collapse Principle established in Part III makes the relationship precise: the observable mixing functional $\Lambda(m)$ is bounded below by a monotone function of the projected sectional curvature $\kappa_\pi(m)$ of the admissibility manifold. Collapse measurements are therefore lower bounds on hidden curvature. They are data from which the hidden geometry can be partially reconstructed. The inverse problem of semantic collapse is: given a field of collapse measurements, recover the admissibility geometry that produced them.

Chapter 2

Existing Frameworks and Their Limits

2.1 Embedding-Based Accounts

The standard picture of meaning in high-dimensional representational systems is geometric: tokens are points, semantic similarity is proximity, and meaning relations are encoded in the local structure of neighborhoods. This picture has proved remarkably productive. It explains fluency, supports generalization, and underlies the practical success of large-scale language models. It also has a structural limitation that becomes visible precisely where the models fail.

The limitation is that geometric proximity is a property of representations, not of the structures representations are meant to encode. Two tokens that are proximate in embedding space may lead to radically different futures under the constraint structure that governs the domain. A representation that identifies them because they are geometrically close conflates things that must remain distinct for inference to work. Embedding-based accounts explain collapse by appealing to the properties of embeddings—the geometry does not separate the relevant classes—without asking why the geometry fails to separate them or what it would mean for it to succeed.

The answer requires going below the level of representations to the constraint structure they project from.

2.2 Modal Logic Approaches

The modal logic response to collapse is more sophisticated. It identifies, correctly, that the distinctions being lost are operator-sensitive: they involve the difference between what is true and what is necessary, what is known and what is the case, what an agent does and what merely happens. It proposes to restore these distinctions by enriching the representational apparatus with accessibility relations, modal bases, and operator-sensitive constraints.

This diagnosis is accurate. The engineering proposal, however, remains within the representational paradigm. Modal Proofing Kernels, typed embeddings, and frame constraints are mechanisms for ensuring that a richer set of distinctions survives compression. They add structure on top of representations rather than deriving representations from structure. The modal operators are labels applied to tokens rather than descriptions of the accessibility geometry that generates meaning. As a consequence, the framework faces a regress: the labels must be sourced from somewhere, and if they are learned from data, they inherit the very ambiguity the framework is designed to cure.

2.3 The Regress Problem

Both embedding-based and modal logic approaches share a hidden assumption: that the categories to be preserved can be identified prior to, and independently of, the construction of the geometry that is supposed to preserve them. The categories are treated as given, and the task is to build a representation that keeps them separate.

But where do the categories come from? If modal labels are learned from distributional data, they inherit the distributional ambiguity of the data. If they are imposed externally, the system presupposes a prior distinction-preserving mechanism—the very thing it is trying to construct. This regress is not a technical problem soluble within the existing paradigm. It is a symptom that the paradigm is asking the wrong question.

The right question is not how to preserve given categories in a representation. The right question is what generates the categories in the first place. The answer, this monograph argues, is the admissibility structure: the field of constraints governing which transitions are and are not permissible. Categories are not primitive. They are descriptions of regions in the admissibility geometry that have distinct accessibility profiles. Representations that preserve categories must preserve accessibility profiles. Representations that collapse categories are destroying accessibility structure.

2.4 The Inversion: Constraints Before Representations

The construction that follows inverts the standard order. Admissibility structure is primary. Representations are projections of that structure. Collapse is what happens when the projection destroys curvature the structure contains. The task is not to add constraints to representations. The task is to understand what constraint geometry generates representations, to measure how much of that geometry survives projection, and to reconstruct it from the collapse patterns that the projection produces.

This inversion has practical consequences. A system designed constraint-first does not generate outputs and filter them for admissibility. It generates within the admissible region from the beginning. The admissibility check is not a downstream repair. It is the primitive computational operation from which generation is derived.

Part II

The Geometry

Chapter 3

The Admissibility Manifold

3.1 Semantic Trajectories as the Primary Objects

Rather than tokens, propositions, or embedding vectors, the primary objects of the theory developed here are trajectories through a space of admissible transitions. A semantic state is not a point endowed with intrinsic properties. It is a position within a structured space of possible continuations, a location in a field that determines what moves forward from it are coherent, what moves are forbidden, and what moves are merely possible.

This shift in primary objects is not terminological. It changes the ontological order of the theory. States have the properties they have because of the transitions they admit. Similarity between states is a function of similarity in their accessibility profiles. The geometry of the state space is induced by the structure of admissible transitions. Nothing in this construction assumes that states have intrinsic properties independent of the constraint field. The field is prior.

Let X denote the space of all such semantic positions. The topology of X is determined by the admissibility relation: two positions are nearby if their accessible futures are similar. The manifold structure, developed in the following sections, is induced by requiring that this similarity vary smoothly across X.

3.2 The Admissible Continuation Space

For each $x \in X$, let $\mathcal{A}(x)$ denote the space of admissible continuations from x, equipped with a conditional probability density $p(y \mid x)$ over $y \in \mathcal{A}(x)$. The density p encodes not merely which continuations are possible but how strongly each is supported by the constraint structure at x. Continuations with high density are those for which the constraints bearing on x provide strong positive support. Those with low density are continuations that are technically permissible but carry little constraint force. Those outside $\mathcal{A}(x)$ are excluded entirely.

Several remarks about the nature of $p(y \mid x)$ are in order. First, it is not learned from data in the sense that a language model learns a next-token distribution. It is the object from which such learning is derived. The density p encodes the underlying constraint structure of the domain, and any learned distribution is an approximation to it. Second, p need not be uniform over $\mathcal{A}(x)$. The admissibility structure assigns different weights to different continuations, and this differential weighting is the source of the metric structure developed in the following chapter. Third, p can vary smoothly with x even when $\mathcal{A}(x)$ itself changes topology—when bottlenecks open or close—provided the transitions in density are continuous.

3.3 Local Charts from Admissibility Profiles

The manifold structure of X is not assumed. It is derived from the admissibility distribution. Define the *admissibility map*:

$$\Psi : X \to \mathcal{P}(\Omega), \qquad \Psi(x) = p(\cdot \mid x),$$

where $\mathcal{P}(\Omega)$ denotes the space of probability measures on the continuation space Ω, equipped with the topology of weak convergence. The map Ψ sends each semantic position to its distribution over admissible continuations.

Proposition 3.1. *If $p(y \mid x)$ depends smoothly on x in the sense that the map $x \mapsto \log p(y \mid x)$ is C^∞ for $p(\cdot \mid x)$-almost every y, and if Ψ is injective, then X is an immersed submanifold of $\mathcal{P}(\Omega)$, and the Fisher information metric on $\mathcal{P}(\Omega)$ pulls back to a smooth Riemannian metric on X.*

Proof. Smooth dependence of $\log p(y \mid x)$ on x implies that the map $\Psi : X \to \mathcal{P}(\Omega)$ is smooth in the sense of the infinite-dimensional manifold structure on $\mathcal{P}(\Omega)$ given by the Fisher-Rao metric. Injectivity of Ψ means that distinct semantic positions have distinct admissibility distributions: the map separates points. A smooth injective map with injective differential is an immersion. The Fisher information metric on $\mathcal{P}(\Omega)$, given by $G(u, v) = \mathbb{E}_p[uv/p^2]$ for tangent vectors u, v in the $L^2(p)$ sense, pulls back along Ψ to the metric $g_{ij}(x) = \mathbb{E}_{y \sim p(\cdot|x)}[\partial_i \log p \cdot \partial_j \log p]$, which is the Fisher information metric of Definition (7.1). Smoothness of the pullback metric follows from smoothness of Ψ. $\square$

The injectivity assumption in Proposition 3.1 has a clear semantic content: two semantic positions are distinct if and only if they differ in at least one admissible continuation. Positions with identical admissibility distributions are semantically indistinguishable. If injectivity fails, X should be replaced by the quotient $X/\sim$, where $x \sim x'$ iff $p(\cdot \mid x) = p(\cdot \mid x')$.

The proposition establishes that the manifold structure of X is inherited from the geometry of probability measures via Ψ. The admissibility map is not merely a convenient parameterization; it is the geometric embedding of the semantic space into the universal statistical manifold $\mathcal{P}(\Omega)$.

3.4 Regularity Conditions

Three regularity conditions on $p(y \mid x)$ are required for the construction to be well-behaved. Each condition has a semantic interpretation that clarifies what it assumes about the constraint structure.

The first condition is smooth variation of $\mathcal{A}(x)$ with x: as the semantic position changes, the accessible futures change continuously. This means that the constraint structure has no sharp discontinuities—no locations where a small change in position produces a catastrophic change in what is admissible. Domains with genuine phase boundaries, where the constraint structure changes sharply, require a separate analysis. The smooth-variation condition holds in the generic case.

The second condition is differentiability of $p(y \mid x)$ in x: the score function $\partial_x \log p(y \mid x)$ exists and is square-integrable with respect to $p(\cdot \mid x)$ for each x. This is the standard regularity condition for Fisher information to be well-defined. It means that the admissibility pressure on each continuation varies smoothly with position.

The third condition is integrability: $\int_{\mathcal{A}(x)} \|\partial_x \log p(y \mid x)\|^2 p(y \mid x)\, dy < \infty$ for each x. This ensures that the Fisher information is finite and the metric can be computed.

3.5 Bottlenecks, Constraints, and Degeneracies

The admissibility manifold X is not homogeneous. Its structure encodes the full topology of the constraint space. Regions where $\mathcal{A}(x)$ is large and the density $p(\cdot \mid x)$

is spread broadly correspond to positions of high semantic flexibility: many coherent continuations are available. Regions where $\mathcal{A}(x)$ contracts sharply correspond to semantic bottlenecks: very few continuations remain admissible, and the constraint structure bears heavily on which way forward is available.

These structural features are encoded in the Fisher metric developed in the next chapter. High curvature in the metric corresponds to regions where the accessibility structure changes rapidly across neighboring positions. Low curvature corresponds to regions where accessibility is stable. The bottlenecks and constraint boundaries of the domain appear as curvature concentrations in the admissibility geometry.

Chapter 4

Tangent Spaces and Admissibility Flows

Before introducing the Fisher metric, we develop the differential structure of the admissibility manifold directly from the admissibility distributions. This chapter establishes the tangent bundle of X as a space of score-compatible perturbations and derives the inner product on each tangent space. The Fisher metric then emerges naturally as the unique metric compatible with this structure, rather than being defined externally.

4.1 Statistical Tangent Vectors

At each point $x \in X$, the tangent space $T_x X$ consists of infinitesimal deformations of the admissibility profile $p(\cdot \mid x)$. A deformation of $p(\cdot \mid x)$ in direction u must preserve the normalization constraint $\int_\Omega p(y \mid x)\, dy = 1$. Differentiating with respect to a curve $x(t)$ through x gives:

$$\frac{d}{dt}\bigg|_{t=0} \int_\Omega p(y \mid x(t))\, dy = \int_\Omega \dot{x}^i(0)\, \partial_i p(y \mid x)\, dy = 0.$$

Writing $u(y) = \dot{x}^i(0)\, \partial_i \log p(y \mid x)$ (the score in direction $\dot{x}$), this normalization constraint becomes:

$$\int_\Omega u(y)\, p(y \mid x)\, dy = 0.$$

Definition 4.1 (Statistical Tangent Space)**.** The statistical tangent space to X at x is the $L^2(p(\cdot \mid x))$-closure:

$$T_x X = \left\{ u \in L^2(\Omega, p(\cdot \mid x)) : \int_\Omega u(y)\, p(y \mid x)\, dy = 0 \right\}. \tag{4.1}$$

Elements of $T_x X$ are called *admissibility flows* at x.

The zero-mean condition in (4.1) is the infinitesimal form of normalization preservation. An admissibility flow u specifies how the distribution over accessible futures changes as the semantic position moves in X: positive values of $u(y)$ indicate that continuation y becomes more accessible, negative values indicate it becomes less accessible, and the constraint $\int u\, p\, dy = 0$ ensures that accessibility is redistributed rather than created or destroyed.

4.2 Score Functions as Natural Coordinates

The partial derivatives of the log-density provide a canonical basis for the statistical tangent space.

Proposition 4.2 (Score Functions Span the Tangent Space)**.** *Under the regularity conditions of Chapter 3, the score functions:*

$$u_i(y) = \partial_i \log p(y \mid x), \qquad i = 1, \dots, \dim X, \tag{4.2}$$

lie in $T_x X$ and, if $\Psi : X \to \mathcal{P}(\Omega)$ is an immersion (Proposition 3.1), they span the image of $d\Psi_x$ in $T_{\Psi(x)}\mathcal{P}(\Omega)$.

Proof. First, $u_i \in T_x X$: by the score equation, $\int \partial_i \log p(y \mid x)\, p(y \mid x)\, dy = \partial_i \int p(y \mid x)\, dy = 0$. So each u_i has zero $p(\cdot \mid x)$-mean.

Second, these are the images of the coordinate vectors $\partial_i \in T_x X$ under $d\Psi_x$: by the chain rule, $\frac{d}{dt}\big|_0 \Psi(x + te_i) = \partial_i p(\cdot \mid x) = p(\cdot \mid x) \cdot \partial_i \log p(\cdot \mid x) = p \cdot u_i$. Thus $d\Psi_x(\partial_i) = p \cdot u_i$ in L^2, and the u_i span the image of $d\Psi_x$. □

Proposition 4.2 identifies the score functions as the natural coordinate vectors on X viewed as a submanifold of $\mathcal{P}(\Omega)$. The score $u_i(y) = \partial_i \log p(y \mid x)$ measures the sensitivity of the log-accessibility of continuation y to a displacement in direction i in semantic space.

4.3 The Fisher Inner Product

The inner product on each tangent space is determined by requiring that the metric measure local statistical distinguishability.

Definition 4.3 (Fisher Inner Product)**.** The Fisher inner product on $T_x X$ is:

$$g_x(u,v) = \int_\Omega u(y)\, v(y)\, p(y \mid x)\, dy = \mathbb{E}_{y\sim p(\cdot|x)}[u(y)\, v(y)]. \tag{4.3}$$

Lemma 4.4 (Properties of the Fisher Inner Product)**.** *The bilinear form* (4.3) *is:*

(i) Bilinear and symmetric*: immediate from the definition.*

(ii) Positive semi-definite*: $g_x(u,u) = \mathbb{E}[u^2] \geq 0$ with equality iff $u = 0$ $p(\cdot \mid x)$-a.e.*

(iii) Non-degenerate on the tangent space*: if $g_x(u,v) = 0$ for all $v \in T_x X$, then $u = 0$ $p(\cdot \mid x)$-a.e., because $T_x X$ is a dense subspace of $L^2(\Omega, p(\cdot \mid x))$ and separates points.*

When evaluated on the score functions (4.2), the Fisher inner product recovers exactly the Fisher information metric:

$$g_x(u_i, u_j) = \mathbb{E}_{y\sim p(\cdot|x)}[\partial_i \log p(y \mid x) \cdot \partial_j \log p(y \mid x)] = g_{ij}(x),$$

as in equation (7.1). The Fisher metric is therefore not introduced by definition but derived as the coordinate representation of the inner product on the statistical tangent space in the natural score basis.

4.4 Uniqueness of the Fisher Metric

The Fisher inner product is the unique Riemannian metric on X compatible with infinitesimal statistical distinguishability. We make this precise through the following axiomatic characterization, which is the classical result of Chentsov adapted to the admissibility setting [2].

Theorem 4.5 (Uniqueness of the Fisher Metric)**.** *Let h be any Riemannian metric on X such that:*

(i) Statistical invariance*: h is invariant under sufficient statistics, i.e., if $T : \Omega \to \Omega'$ is a sufficient statistic for the family $\{p(\cdot \mid x)\}$, then the pushforward metric under T equals h.*

(ii) Positive definiteness*: $h_x(u,u) > 0$ for all non-zero $u \in T_x X$.*

Then $h = c\, g$ for some positive constant c. Up to scale, the Fisher metric is the unique metric satisfying statistical invariance.

The statistical invariance condition means that if two parameterizations of the admissibility family lead to the same distributions, then any sufficient statistic that mediates between them should not change measured distances. This is the information-theoretic content of the uniqueness: the Fisher metric measures information, and information is invariant under sufficient compression.

4.5 Semantic Interpretation of Admissibility Flows

A tangent vector $u \in T_x X$ admits a direct semantic interpretation. Writing $u(y) = \partial_i \log p(y \mid x)\, v^i$ for some velocity $v = v^i \partial_i \in T_x X$ (using the score basis), we have:

$$u(y) = \frac{d}{d\epsilon}\bigg|_0 \log p(y \mid x + \epsilon v).$$

This is the rate of change of the log-accessibility of continuation y as the semantic position moves in direction v. Continuations with large positive $u(y)$ are those that become significantly more accessible; those with large negative $u(y)$ become less accessible. The admissibility flow u therefore specifies the direction in which the accessible future is deforming as the semantic trajectory moves through X.

The Fisher metric $g_x(u, u) = \mathbb{E}[u^2]$ measures the mean squared rate of change of log-accessibility: how rapidly the distribution over accessible futures is deforming. Large values indicate semantically loaded directions where small positional changes strongly reshape the future. Small values indicate semantically redundant directions where the future is stable. This is the intrinsic content of the Fisher metric as a measure of semantic distinguishability.

Chapter 5

Admissibility Geodesics

With the tangent space structure in place, we can define geodesics on the admissibility manifold. Geodesics are the paths of minimal semantic deformation: they represent the most efficient routes between semantic positions, deforming the admissibility structure as little as possible along the way. Before discussing curvature in Chapter 5, we need geodesics to give curvature something to act on: curvature measures how geodesics diverge from the flat case.

5.1 The Geodesic Energy Functional

Let $\gamma : [0,1] \to X$ be a smooth curve connecting $\gamma(0) = x_0$ to $\gamma(1) = x_1$. The length of γ in the Fisher metric is:

$$L[\gamma] = \int_0^1 \sqrt{g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))}\, dt,$$

and the energy functional is:

$$E[\gamma] = \frac{1}{2} \int_0^1 g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))\, dt. \tag{5.1}$$

By the Cauchy-Schwarz inequality, $L[\gamma]^2 \leq 2E[\gamma]$ with equality when $|\dot{\gamma}|$ is constant. Therefore minimizing energy among constant-speed reparameterizations is equivalent to minimizing length.

5.2 The Geodesic Equation

Theorem 5.1 (Admissibility Geodesic Equation)**.** *Critical points of the energy functional* (5.1) *satisfy the Euler-Lagrange equation:*

$$\ddot{x}^k + \Gamma^k_{ij}(x)\,\dot{x}^i\dot{x}^j = 0, \tag{5.2}$$

where Γ^k_{ij} *are the Christoffel symbols of the Fisher metric:*

$$\Gamma^k_{ij} = \frac{1}{2}g^{kl}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right). \tag{5.3}$$

Solutions to (5.2) *are* admissibility geodesics*.*

Proof. Standard variational calculus. Consider a variation $\gamma_\epsilon(t) = \gamma(t) + \epsilon\eta(t)$ with $\eta(0) = \eta(1) = 0$. Differentiating $E[\gamma_\epsilon]$ with respect to ϵ at $\epsilon = 0$ and setting equal to zero:

$$\left.\frac{d}{d\epsilon}\right|_0 E[\gamma_\epsilon] = \int_0^1 \left[g_{ij}\ddot{\gamma}^i\eta^j + \frac{1}{2}\partial_k g_{ij}\dot{\gamma}^i\dot{\gamma}^j\eta^k\right] dt = 0,$$

integrating the first term by parts. Since η is arbitrary, the integrand vanishes pointwise, giving:

$$g_{kj}\ddot{\gamma}^j + \partial_k g_{ij}\dot{\gamma}^i\dot{\gamma}^j - \frac{1}{2}\partial_k g_{ij}\dot{\gamma}^i\dot{\gamma}^j = 0.$$

Contracting with g^{kl} and symmetrizing in i, j yields (5.2) with (5.3). □

5.3 Fisher Geodesic Distance

The geodesic distance between two semantic positions is the infimum of lengths over all smooth paths:

$$d_F(x_1, x_2) = \inf_{\gamma:\gamma(0)=x_1,\,\gamma(1)=x_2} L[\gamma], \tag{5.4}$$

where the infimum is over smooth curves in X. When a minimizing geodesic exists (which it does by the Hopf-Rinow theorem whenever (X, g) is complete), d_F is realized by a solution to (5.2). The Fisher distance $d_F(x_1, x_2)$ measures the minimum semantic deformation required to move from admissibility profile $p(\cdot \mid x_1)$ to $p(\cdot \mid x_2)$: it is the information-geometric cost of the transition.

The distance d_F is the quantity that appears in the curvature-divergence inequality (10.2) of Chapter 6. That inequality can now be restated with the geodesic definition in place: the Jensen-Shannon divergence between nearby admissibility distributions grows at most quadratically in the geodesic distance, with the coefficient of growth given by the Lipschitz constant of the admissibility density.

5.4 Semantic Interpretation: Geodesics as Minimal Deformation

An admissibility geodesic $\gamma : [0,1] \to X$ is the path of minimal semantic deformation from x_0 to x_1: it traverses the least-curved route through the admissibility manifold, redistributing the accessible future as smoothly as possible along the way.

In flat regions of X (where the curvature R_g vanishes), geodesics are straight lines in the statistical sense: the admissibility profile deforms linearly from $p(\cdot \mid x_0)$ to $p(\cdot \mid x_1)$. In curved regions, geodesics are deflected by the curvature. The direction of deflection is toward regions of higher admissibility density—toward the semantic attractors where the accessible future is richest. This is the information-geometric analogue of gravitational lensing: curvature bends the minimal-deformation paths toward regions of concentrated constraint.

5.5 Bottleneck Deflection and Escape Velocity

A semantic bottleneck at $x_b \in X$ is a region where $\mathcal{A}(x_b)$ is small and the admissibility density $p(\cdot \mid x_b)$ is concentrated on a small number of continuations. In the Fisher metric, bottlenecks correspond to regions where the metric degenerates in the directions of collapsed futures: the accessible-future cone contracts, and geodesics approaching x_b from the side must deflect to avoid the degenerate directions.

The geometry near a bottleneck is analogous to the geometry near a saddle point in Riemannian geometry: geodesics arriving along the wide direction of the saddle pass through relatively freely, while those arriving along the narrow direction are deflected or blocked. The bottleneck creates an effective potential in the geodesic equation, which can be made precise via the Jacobi equation (the linearization of (5.2)) and the associated conjugate point theory.

Chapter 6

Volume Growth and Accessibility Geometry

Curvature in a Riemannian manifold controls the rate at which volumes of geodesic balls grow with radius. In the admissibility manifold, this volume growth has a direct semantic interpretation: it measures how rapidly the accessible future expands as the semantic position moves away from a given point. This chapter establishes the Bishop-Gromov comparison theorem for the admissibility manifold and uses it to formalize the bottleneck language that has appeared throughout the preceding chapters.

6.1 Fisher Balls and Accessibility Volume

Definition 6.1 (Fisher Ball)**.** The Fisher ball of radius r centered at $x \in X$ is:

$$B_r(x) = \{y \in X : d_F(x, y) < r\}, \tag{6.1}$$

and its volume with respect to the Riemannian volume form $d\mu_g$ of (X, g) is:

$$V(x, r) = \mu_g(B_r(x)) = \int_{B_r(x)} d\mu_g. \tag{6.2}$$

The Fisher ball $B_r(x)$ is the set of semantic positions reachable from x by a geodesic of Fisher length at most r. In the admissibility interpretation, it is the set of semantic positions whose admissibility profiles differ from $p(\cdot \mid x)$ by a KL divergence of at most $\frac{1}{2}r^2$ (by Theorem 7.1). The volume $V(x, r)$ measures the size of the admissibility-accessible neighborhood of x at scale r.

6.2 Comparison Geometry: The Bishop-Gromov Theorem

In a Riemannian manifold with Ricci curvature bounded below by a constant K, the volume of geodesic balls satisfies a comparison inequality with the corresponding volume in the space form of constant curvature K. Let $V_K(r)$ denote the volume of a ball of radius r in the n-dimensional space of constant sectional curvature K:

$$V_K(r) = \omega_{n-1} \int_0^r \mathrm{sn}_K(t)^{n-1}\, dt,$$

where ω_{n-1} is the volume of the unit $(n-1)$-sphere and:

$$\mathrm{sn}_K(t) = \begin{cases} K^{-1/2} \sin(K^{1/2} t) & K > 0, \\ t & K = 0, \\ |K|^{-1/2} \sinh(|K|^{1/2} t) & K < 0. \end{cases}$$

Theorem 6.2 (Bishop-Gromov for the Admissibility Manifold)**.** *Let (X, g) be the admissibility manifold with Fisher metric g, and suppose the Ricci curvature satisfies $\mathrm{Ric}_g \geq K$ for some $K \in \mathbb{R}$ and all $x \in X$. Then the ratio:*

$$r \mapsto \frac{V(x,r)}{V_K(r)} \tag{6.3}$$

is monotone non-increasing in r for all $x \in X$.

Proof. This is the classical Bishop-Gromov theorem applied to the Riemannian manifold (X, g) with the Fisher metric. The proof uses the Jacobi field comparison: if $\mathrm{Ric}_g \geq K$, then Jacobi fields in (X, g) grow no faster than Jacobi fields in the comparison space of curvature K. Integrating Jacobi field lengths over the unit sphere in the tangent space gives the volume comparison (6.3). The monotonicity follows from the Riccati comparison for the mean curvature of distance spheres. □

6.3 Curvature and Accessibility Volume

Theorem 6.2 has direct semantic consequences. The monotonicity of $V(x,r)/V_K(r)$ translates as follows.

Corollary 6.3 (Positive Curvature Contracts Accessible Futures). *If* $\mathrm{Ric}_g(x) > 0$, *then* $V(x,r) < V_0(r) = \omega_{n-1}r^n/n$ *for all sufficiently small* $r > 0$*: the Fisher ball at* x *is smaller than the Euclidean ball of the same radius. Semantically, positive Ricci curvature at* x *means that the admissibility-accessible neighborhood of* x *is smaller than it would be in a flat admissibility geometry. The constraint structure at* x *concentrates accessible futures into a smaller volume.*

Corollary 6.4 (Negative Curvature Expands Accessible Futures). *If* $\mathrm{Ric}_g(x) < 0$, *then* $V(x,r) > V_0(r)$ *for small* r*: the Fisher ball is larger than the Euclidean ball. Negative Ricci curvature at* x *means that the accessible neighborhood expands faster than it would in flat space. The constraint structure at* x *is locally divergent: small departures from* x *lead to rapidly diversifying accessible futures.*

6.4 Semantic Bottlenecks as Volume Suppression

The bottleneck language used throughout this monograph can now be formalized precisely.

Definition 6.5 (Semantic Bottleneck). A point $x \in X$ is an (ϵ, r)-*bottleneck* if:

$$V(x,r) \le \epsilon\, V_0(r), \tag{6.4}$$

i.e., the Fisher ball of radius r at x has volume at most ϵ times the Euclidean ball of the same radius. An (ϵ, r)-bottleneck is a region where accessible futures are strongly suppressed relative to the flat baseline.

By Corollary 6.3, points of large positive Ricci curvature are bottlenecks. The severity of the bottleneck is controlled by the magnitude of Ric_g: the larger the Ricci curvature, the smaller $V(x,r)/V_0(r)$, and the more tightly the accessible futures are constrained.

The bottleneck structure connects to the projection analysis of Chapter 5. A projection $\pi : X \to M$ that maps a bottleneck region to a low-dimensional region in M must identify many admissibility-distinct points in the fiber over that region: the volume suppression at the bottleneck forces the fiber to be large. By (8.6), large fibers in curved regions produce large projected sectional curvature κ_π. The Projection-Collapse Principle then forces large Λ: bottlenecks are the primary sources of observable collapse.

Chapter 7

The Fisher Information Metric

7.1 Why the Hessian Fails

The natural first attempt at a metric on X would be to construct a scalar functional $\Phi : X \to \mathbb{R}$ measuring the total accessibility at each position—the log-volume of the admissible continuation space, for instance—and to define the metric as its Hessian:

$$g_{ij}(x) = \partial_i \partial_j \Phi(x).$$

This construction has two advantages. It is familiar from optimization and from the differential geometry of level sets. And it produces a metric that degenerates at the critical points of Φ, which might seem to correspond to semantic bottlenecks.

The Hessian construction fails for a precise reason. Collapse events occur at saddle points, ridges, and near-degenerate regions of Φ—exactly the locations where the Hessian is indefinite or vanishing. At a saddle point, the Hessian has both positive and negative eigenvalues: it is not a metric at all. At a flat region near a bottleneck, the Hessian is nearly zero and provides no geometric information. The most semantically interesting locations in X—the places where the constraint structure changes rapidly and compression destroys the most information—are precisely the locations where the Hessian fails.

7.2 The Fisher Information Metric

The correct metric is the Fisher information metric induced by the conditional density $p(y \mid x)$. Define:

$$g_{ij}(x) = \mathbb{E}_{y \sim p(\cdot|x)} \left[\partial_i \log p(y \mid x) \cdot \partial_j \log p(y \mid x) \right]. \tag{7.1}$$

Equivalently, under the regularity conditions of Chapter 3:

$$g_{ij}(x) = -\mathbb{E}_{y\sim p(\cdot|x)}\left[\partial_i \partial_j \log p(y \mid x)\right].$$

The equivalence between these two expressions is standard in information geometry and follows from differentiating the normalization condition $\int p(y \mid x)\, dy = 1$ under the integral sign. The second expression is sometimes more tractable for computation; the first is more directly interpretable.

7.3 Derivation from Infinitesimal Admissibility Divergence

The Fisher metric admits a derivation that grounds it geometrically in the distinguishability of nearby semantic positions. Consider two nearby states x and $x + dx$ in X. The KL divergence from $p(\cdot \mid x)$ to $p(\cdot \mid x + dx)$ measures the information cost of replacing the admissibility distribution at x with that at $x + dx$:

$$D_{\mathrm{KL}}(p(\cdot \mid x) \parallel p(\cdot \mid x + dx)) = \int_\Omega p(y \mid x) \log \frac{p(y \mid x)}{p(y \mid x + dx)}\, dy.$$

Theorem 7.1. *Under the regularity conditions of Section 3.3, the KL divergence between nearby admissibility distributions satisfies:*

$$D_{\mathrm{KL}}(p(\cdot \mid x) \parallel p(\cdot \mid x + dx)) = \frac{1}{2} g_{ij}(x)\, dx^i\, dx^j + O(\|dx\|^3), \tag{7.2}$$

where g_{ij} is the Fisher information metric (7.1). *Moreover, g is the unique Riemannian metric on X whose quadratic form equals the second-order approximation to admissibility divergence.*

Proof. Expand $\log p(y \mid x + dx)$ in Taylor series:

$$\log p(y \mid x+dx) = \log p(y \mid x) + \partial_i \log p(y \mid x)\, dx^i + \tfrac{1}{2}\partial_i\partial_j \log p(y \mid x)\, dx^i dx^j + O(\|dx\|^3).$$

Substituting into the KL divergence:

$$D_{\mathrm{KL}}(p(\cdot \mid x) \parallel p(\cdot \mid x + dx)) = -\int p(y \mid x)\left[\partial_i \log p(y \mid x)\, dx^i + \tfrac{1}{2}\partial_i\partial_j \log p(y \mid x)\, dx^i dx^j\right] dy + O(\|$$

The first-order term vanishes by the score equation: differentiating $\int p(y \mid x)\, dy = 1$ gives $\int p(y \mid x)\, \partial_i \log p(y \mid x)\, dy = 0$.

For the second-order term, differentiate the score equation with respect to x^j:

$$\int \partial_j p(y \mid x) \cdot \partial_i \log p(y \mid x)\, dy + \int p(y \mid x) \cdot \partial_i \partial_j \log p(y \mid x)\, dy = 0.$$

The first integral equals $\int p(y \mid x)\, (\partial_j \log p)(\partial_i \log p)\, dy = g_{ij}(x)$, so:

$$\int p(y \mid x)\, \partial_i \partial_j \log p(y \mid x)\, dy = -g_{ij}(x).$$

Therefore the second-order term contributes $\frac{1}{2} g_{ij}(x)\, dx^i dx^j$, which establishes (7.2). Uniqueness follows because the coefficients of $dx^i dx^j$ in the Taylor expansion of a fixed divergence are uniquely determined, forcing $h_{ij} = g_{ij}$ for any metric h with the same second-order approximation property. $\square$

Theorem 7.1 establishes that the Fisher metric is the infinitesimal form of KL divergence between admissibility distributions. Two semantic positions at Fisher distance ϵ have admissibility distributions that differ by approximately $\frac{1}{2}\epsilon^2$ in KL divergence. The metric measures how distinguishable nearby positions are by their accessible futures.

7.4 Positive Semi-Definiteness and Its Semantic Meaning

The Fisher metric (7.1) is positive semi-definite at every $x \in X$ by construction. The outer product $\partial_i \log p \cdot \partial_j \log p$ is a rank-one positive semi-definite matrix at each y, and expectation preserves positive semi-definiteness. No convexity assumption on Φ is required.

The metric degenerates—becomes rank-deficient—precisely when $p(\cdot \mid x)$ is insensitive to perturbations in the direction indexed by i. That is: $g_{ii}(x) = 0$ if and only if $\partial_i \log p(y \mid x) = 0$ for $p(\cdot \mid x)$-almost every y, meaning that moving in direction i does not change the accessible futures. Degeneracy is semantically meaningful. It identifies the directions in X along which the constraint structure is invariant: moving in a degenerate direction does not change what is admissible, and no geometric resolution is lost by collapsing such directions in a representation.

Directions of large metric weight are the semantically loaded directions: small perturbations along them strongly reshape $p(\cdot \mid x)$, changing which futures are accessible. These are the directions that representations must preserve to maintain semantic fidelity. Directions of small metric weight are the semantically redundant directions: the accessibility structure is stable under perturbation, and compression along these directions loses little.

7.5 Connection to Information Geometry

The admissibility manifold (X, g) is a statistical manifold in the sense of Amari: a smooth manifold whose points parameterize a family of probability distributions, equipped with the Fisher information metric [2]. The specific interpretation here differs from the standard statistical setting. The parameter space is not a space of statistical models but a space of semantic positions. The family of distributions is the family of admissibility densities $\{p(\cdot \mid x) : x \in X\}$. The Fisher metric measures not statistical distinguishability in the standard sense but semantic distinguishability: the degree to which different positions lead to different accessible futures.

This reinterpretation carries consequences. The Cramér-Rao inequality, in this setting, becomes a lower bound on how precisely a representation can resolve distinct admissibility structures. Any projection $\pi : X \to M$ that attempts to encode the semantic position x as a point $m = \pi(x)$ in a lower-dimensional space M loses information in proportion to the curvature of (X, g) in the directions collapsed by the projection. The Cramér-Rao bound is not a statistical limitation. It is a geometric limitation on representational fidelity.

7.6 Curvature of the Admissibility Manifold

The Riemann curvature tensor R_g of the Fisher metric measures how the local admissibility structure twists and turns across neighboring regions of X. High curvature at x means that nearby points have substantially different admissible future cones: small displacements produce large changes in which continuations are available. Flat regions—where curvature is low—mean that the accessibility structure varies slowly and projections can represent it faithfully.

The curvature identifies, in geometric terms, the locations where any compression will lose information: where the admissibility geometry is curved, a projection that identifies nearby points will necessarily confuse trajectories that lead to different futures. Flat regions project cleanly. Curved regions project poorly. Semantic collapse is the observable consequence of projecting curved admissibility geometry onto a flatter representational space.

Chapter 8

Projections and the Vertical-Horizontal Decomposition

8.1 Representations as Projections

A representational system is a smooth map $\pi : X \to M$ from the admissibility manifold to a lower-dimensional representational manifold M. The map π compresses X by identifying points that share a representation: for each $m \in M$, the fiber $\pi^{-1}(m)$ is the set of all admissibility states that project to m. Neural encoders, embedding maps, summary statistics, and symbolic abstractions are all instances of this compression. What distinguishes them is not the conceptual form of the operation but the specific properties of π: how much curvature it destroys, which directions it collapses, and how the fiber structure relates to the admissibility geometry.

The faithfulness of a representation is determined by how much of the geometric structure of (X, g) survives the map π. A representation is faithful in direction $v \in T_xX$ if the differential $d\pi_x$ is an isometry in that direction: if it maps v to a vector in $T_{\pi(x)}M$ of the same length. It is unfaithful if $d\pi_x(v) = 0$: if direction v is invisible to π. Collapse occurs when semantically loaded directions—directions of high metric weight in g—are made invisible by the projection.

8.2 The Vertical-Horizontal Decomposition

At each $x \in X$, the tangent space T_xX decomposes canonically into two orthogonal subspaces:

$$T_xX = V_x \oplus H_x, \tag{8.1}$$

where $V_x = \ker(d\pi_x)$ is the *vertical bundle*: the subspace of tangent directions that are invisible to π. These are the directions that the representation cannot see. The subspace $H_x = V_x^\perp$ is the *horizontal bundle*: the g-orthogonal complement of V_x, consisting of the directions that the projection faithfully transmits.

The decomposition (8.1) is the geometric formalization of the distinction between what a representation can see and what it cannot. Horizontal directions are preserved. Vertical directions are destroyed. What matters for collapse is neither pure vertical curvature nor pure horizontal curvature. Pure vertical curvature—curvature entirely within V_x—is hidden by π and averaged over within each fiber, but because it does not couple to horizontal directions, it does not introduce systematic errors at the representational level. Pure horizontal curvature—curvature entirely within H_x—is preserved by π and appears faithfully in M. It is the *mixed* curvature, the sectional curvature $K_g(v,h)$ for $v \in V_x$ and $h \in H_x$, that is dangerous. Mixed curvature couples what is hidden to what is observable. When it is present, the fiber contains points that look identical in M but lead to different observable behaviors depending on which direction they are approached from. This is the geometric source of collapse.

8.3 Projected Sectional Curvature

We define the projected sectional curvature at $m \in M$ as the integrated mixed sectional curvature over the fiber $\pi^{-1}(m)$:

$$\kappa_\pi(m) = \int_{\pi^{-1}(m)} \sum_{\substack{v \in V_x,\, h \in H_x \\ |v|=|h|=1}} \left| K_g(v,h) \right| d\mu_m(x), \tag{8.2}$$

where μ_m is the measure on the fiber induced by the volume form of g and the sum is taken over an orthonormal frame. This is the curvature that projection destroys: the coupling between the hidden directions and the observable directions. Only mixed sectional curvature contributes to the leakage that collapse diagnostics measure. Pure vertical and pure horizontal curvature do not.

The projected sectional curvature κ_π is a refined measure. It isolates, within the full Riemann tensor R_g, exactly the components that a neighborhood-based statistic computed in M can detect. This refinement is not merely technical. It has direct semantic content: $\kappa_\pi(m)$ measures how much of the admissibility structure at m the representation has destroyed in a way that creates observable confusion.

8.4 The O'Neill Tensors and Collapse-Driving Curvature

The precise relationship between the projection geometry and the mixed sectional curvature is given by O'Neill's fundamental equations for Riemannian submersions [21]. These equations decompose the curvature of (X, g) into contributions from the base, the fibers, and the interaction between them. The interaction term is the collapse-driving curvature.

For a smooth vector field E on X, write $E = E^H + E^V$ for its horizontal and vertical components. Define the *O'Neill A-tensor* and *T-tensor* as:

$$A_E F = (\nabla_{E^H} F^H)^V + (\nabla_{E^H} F^V)^H, \tag{8.3}$$

$$T_E F = (\nabla_{E^V} F^V)^H + (\nabla_{E^V} F^H)^V, \tag{8.4}$$

where ∇ is the Levi-Civita connection of g. The A-tensor measures the non-integrability of the horizontal distribution: $A_h v = (\nabla_h v)^V$ for $h \in H_x$ and $v \in V_x$. The T-tensor measures the second fundamental form of the fibers.

Proposition 8.1 (O'Neill Mixed Curvature Formula)**.** *For $v \in V_x$ and $h \in H_x$ with $|v| = |h| = 1$, the mixed sectional curvature is:*

$$K_g(v, h) = g((\nabla_v A)_h h,\, v) - \|A_h v\|^2 + g(T_v h,\, T_v h) - g(T_v v,\, T_h h) + \cdots \tag{8.5}$$

where the dominant collapse-driving term is $\|A_h v\|^2$, the squared norm of the A-tensor acting on the mixed pair (h, v).

The significance of Proposition 8.1 is that it identifies the A-tensor as the primary source of mixed curvature. The quantity $A_h v = (\nabla_h v)^V$ measures how much a horizontal displacement h rotates a vertical direction v: if horizontal motion through the fiber changes which vertical directions are present, then the fiber structure is non-trivially coupled to the base, and projection destroys information. When $A \equiv 0$, the horizontal distribution is integrable, the projection is locally a product, and no mixed curvature is present. In that case $\kappa_\pi = 0$ and no leakage occurs.

This connects directly to the semantic interpretation: the A-tensor measures how much the admissibility structure varies across nearby fibers. High $\|A_h v\|^2$ at x means that moving horizontally by h—changing the representation—changes which vertical directions carry admissibility information. Projection then cannot faithfully represent the admissibility structure, and collapse is geometrically forced.

The projected sectional curvature (8.2) is therefore controlled by the A-tensor norm integrated over the fiber:

$$\kappa_\pi(m) \asymp \int_{\pi^{-1}(m)} \|A(x)\|^2 \, d\mu_m(x), \tag{8.6}$$

where $\|A(x)\|^2 = \sum_{h,v} |A_h v|^2$ with the sum over an orthonormal frame. This is the form of κ_π that enters the proof of the Projection-Collapse Principle.

8.5 The Fiber as a Hidden Variable

The fiber $\pi^{-1}(m)$ is the set of all admissibility states that project to the same representation m. It is a hidden variable in the strict statistical sense: it affects observable outcomes but is not directly accessible through M. Two trajectories that are representationally identical—that pass through the same point in M—may have very different admissible futures if they come from different points in the fiber.

Collapse diagnostics are measurements that partially identify the fiber structure without directly observing it. The mixing functional $\Lambda(m)$, defined in the next chapter, is a statistic computed in M that measures the degree to which the fiber $\pi^{-1}(m)$ contains admissibility-distinct points. The Projection-Collapse Principle establishes the relationship between this observable and the hidden curvature $\kappa_\pi(m)$. The inverse problem is to recover κ_π from measurements of Λ: to reconstruct the hidden geometry from the observable trace it produces through collapse.

Part III

The Theorem

Chapter 9

The Representational Mixing Functional

9.1 Beyond Operator-Specific Leakage

Several diagnostic frameworks for measuring semantic collapse have been proposed in the literature [13]. These frameworks define cross-type leakage statistics that measure the degree to which semantically distinct operator classes mix within embedding neighborhoods. These statistics are well-motivated and empirically useful. They are, however, specific to particular representational architectures, particular operator taxonomies, and particular notions of semantic type. For a domain-independent theory, a more general object is needed.

The representational mixing functional defined here generalizes these statistics in a way that removes all dependence on embeddings, operator classes, and domain-specific categorizations. It measures, at the level of abstract probability theory, the degree to which a projection fails to separate admissibility classes that should remain distinct. The existing diagnostics are then special cases, and the Projection-Collapse Principle applies to all of them simultaneously.

9.2 Definition of the Mixing Functional

Let $U \subset M$ be a measurable region in the representational manifold, and let $\{U_i\}_{i=1}^{k}$ be a measurable partition of U into admissibility classes—subsets that should, under a faithful representation, remain separated because they correspond to distinct regions of the admissibility manifold. Define the representational mixing functional:

$$\Lambda(U) = D\left(\pi_* \mu_U, \prod_{i=1}^{k} \pi_* \mu_{U_i}\right), \tag{9.1}$$

where D is an information divergence and $\pi_* \mu_{U_i}$ is the pushforward of the admissibility measure restricted to class i.

When the admissibility classes are perfectly separated in M—when the projection faithfully distinguishes them—the joint distribution $\pi_* \mu_U$ factorizes over the partition and $\Lambda(U) = 0$. When the classes mix under projection, the joint distribution diverges from the product distribution and $\Lambda(U) > 0$. The divergence quantifies precisely how much admissibility structure the projection has failed to separate.

The choice of information divergence D is a parameter of the framework. The Kullback-Leibler divergence, Jensen-Shannon divergence, and Rényi divergence each produce a version of the functional with different analytical properties. For the strong form of the Projection-Collapse Principle, we require Jensen-Shannon divergence: it is bounded, symmetric, and metrizes weak convergence—properties that the KL divergence lacks and that the asymptotic equivalence proof requires.

9.3 Existing Diagnostics as Special Cases

The cross-type leakage statistics of [13] are instances of (9.1). Specifically, those statistics arise when U is a k-nearest-neighbor ball around a token t in the embedding space, the partition $\{U_i\}$ is into operator type classes, and D is a normalized mass-leakage ratio. The generalization does not change the diagnostic procedure. It changes its interpretation: the mixing functional Λ is not a statistic internal to the embedding space. It is a measure of admissibility class separation failure under projection, and the same quantity applies to any representational compression of any structured space.

This interpretive shift has a specific consequence. When an existing diagnostic reports high leakage for modal tokens, it is not reporting a property of the embedding. It is reporting a measurement of the projected sectional curvature κ_π of the admissibility manifold, filtered through a particular choice of U, partition, and divergence. The Projection-Collapse Principle makes the relationship between the measurement and the underlying geometry explicit.

9.4 Regularity Conditions on the Functional

Two regularity conditions on Λ are required for the Projection-Collapse Principle. First, continuity with respect to the pushforward measure $\pi_*\mu$: small changes in the distribution of representations produce small changes in the mixing functional. This is satisfied by all standard information divergences under mild conditions on the underlying densities.

Second, measurability of the partition $\{U_i\}$: the admissibility classes can be identified by measurable sets in M. This is a non-trivial condition in practice. It requires that the admissibility classes have sufficiently regular boundaries in the representational space. When the projection π severely distorts the admissibility geometry, the images of admissibility classes may become non-measurable or have fractal boundaries. The regularity condition restricts the analysis to projections that, while potentially destructive of curvature, at least preserve the measurability of class structure.

Chapter 10

The Projection-Collapse Principle

10.1 Fiber Divergence: The Bridge Between Curvature and Leakage

Before stating the main theorem, we establish the intermediate object that connects curvature to leakage. The fiber divergence measures the internal admissibility heterogeneity of each fiber, providing the missing mathematical bridge in the chain $\kappa_\pi \to \Delta_F \to \Lambda$.

Definition 10.1 (Fiber Divergence)**.** Let $\bar{p}_m = \frac{1}{\mu_m(\pi^{-1}(m))} \int_{\pi^{-1}(m)} p(\cdot \mid x)\, d\mu_m(x)$ be the barycenter of the admissibility distributions over the fiber $\pi^{-1}(m)$. Define the *fiber divergence* at $m \in M$ as:

$$\Delta_F(m) = \int_{\pi^{-1}(m)} D_{\mathrm{JS}}(p(\cdot \mid x),\, \bar{p}_m)\, d\mu_m(x), \tag{10.1}$$

where D_{JS} is the Jensen-Shannon divergence.

Lemma 10.2 (Fiber Divergence and Admissibility Uniformity)**.** $\Delta_F(m) = 0$ *if and only if $p(\cdot \mid x) = \bar{p}_m$ for μ_m-almost every $x \in \pi^{-1}(m)$: that is, every point in the fiber has identical admissibility distribution.*

Proof. The Jensen-Shannon divergence $D_{\mathrm{JS}}(p, q) \geq 0$ with equality if and only if $p = q$ as measures. The integral $\Delta_F(m) = \int D_{\mathrm{JS}}(p(\cdot \mid x), \bar{p}_m)\, d\mu_m(x)$ is zero if and only if the integrand is zero μ_m-almost everywhere, which holds if and only if $p(\cdot \mid x) = \bar{p}_m$ for μ_m-a.e. x. □

Lemma 10.2 establishes that $\Delta_F(m) = 0$ is the condition for the fiber to be admissibility-homogeneous: all points in $\pi^{-1}(m)$ lead to the same distribution over accessible futures. When the fiber is admissibility-homogeneous, the projection π

loses no admissibility information at m, and no collapse occurs. $\Delta_F(m) > 0$ is the condition for heterogeneity: different points in the fiber lead to different futures, and the projection conflates them.

10.2 The Curvature-Divergence Inequality

The second step in the proof chain connects curvature to fiber divergence. We show that $\kappa_\pi(m)$ controls $\Delta_F(m)$ via a comparison geometry estimate.

Proposition 10.3 (Curvature-Divergence Inequality)**.** *Let $d_F(x_1, x_2)$ denote the Fisher geodesic distance on (X, g). Under the Lipschitz condition that $p(\cdot \mid x)$ is L_p-Lipschitz in x with respect to d_F, the Jensen-Shannon divergence between admissibility distributions at two points in the same fiber satisfies:*

$$D_{\mathrm{JS}}(p(\cdot \mid x_1),\, p(\cdot \mid x_2)) \le C\, d_F(x_1, x_2)^2, \tag{10.2}$$

for a constant C depending only on L_p. Integrating over the fiber:

$$\Delta_F(m) \le C \int_{\pi^{-1}(m)} d_F(x, \bar{x}_m)^2 \, d\mu_m(x), \tag{10.3}$$

where $\bar{x}_m$ is the fiber barycenter in (X, g).

Proof. The Jensen-Shannon divergence satisfies the bound $D_{\mathrm{JS}}(p, q) \le \|p - q\|_{\mathrm{TV}}$ where $\|\cdot\|_{\mathrm{TV}}$ is the total variation distance, and by Pinsker's inequality $D_{\mathrm{JS}}(p, q) \le \frac{1}{2\ln 2} D_{\mathrm{KL}}(p\|q)$ up to symmetrization. The Lipschitz condition on $p(\cdot \mid x)$ combined with Theorem 7.1 gives:

$$D_{\mathrm{KL}}(p(\cdot \mid x_1) \,\|\, p(\cdot \mid x_2)) \le L_p^2\, d_F(x_1, x_2)^2 + O(d_F^3),$$

which yields (10.2) for small d_F with $C = L_p^2/(2\ln 2)$. The fiber integral (10.3) then follows by integrating over $\pi^{-1}(m)$ and applying the triangle inequality for d_F. □

The right side of (10.3) is a fiber-integrated squared Fisher distance. By comparison geometry—specifically, the Bishop-Gromov volume comparison theorem adapted to the submersion setting—this integral is controlled by the sectional curvature of (X, g) restricted to the fiber. In the regime of bounded curvature, the integral grows with the A-tensor norm via the O'Neill formula (8.5). The complete chain is therefore:

$$\|A\|^2 \quad \longrightarrow \quad \kappa_\pi \quad \longrightarrow \quad d_F(x_1, x_2)^2 \quad \longrightarrow \quad \Delta_F \quad \longrightarrow \quad \Lambda. \tag{10.4}$$

Each arrow corresponds to a proved or bounded inequality. The chain makes explicit what was previously described in prose: collapse is not a coincidence but a geometric consequence, traceable step by step from the non-integrability of the horizontal distribution to the observable mixing functional.

10.3 Statement: Weak Form

Theorem 10.4 (Projection-Collapse Principle, Weak Form)**.** *Let (X, g) be an admissibility manifold equipped with the Fisher information metric* (7.1) *induced by the conditional density $p(y \mid x)$ satisfying the regularity conditions of Chapter* **??**. *Let $\pi : X \to M$ be a smooth submersion with fiber measure μ_m induced by the volume form of g. Let Λ be a representational mixing functional* (9.1) *continuous with respect to the pushforward measure $\pi_*\mu$. Then there exists a monotone non-decreasing function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with $f(0) = 0$ such that*

$$\Lambda(m) \geq f(\kappa_\pi(m)) \tag{10.5}$$

for all $m \in M$. Sufficiently large hidden admissibility curvature forces observable representational mixing.

10.4 Proof of the Weak Form

The argument proceeds in four steps.

Step 1. High projected sectional curvature implies admissibility-distinct points in the fiber. If $\kappa_\pi(m) > 0$, then by definition (8.2) there exists a set of positive measure in $\pi^{-1}(m)$ on which the mixed sectional curvature $|K_g(v, h)|$ is positive. At such points, parallel transport along horizontal curves rotates the vertical bundle, meaning that nearby fibers have different admissibility structures. By the Cramér-Rao argument applied to the Fisher metric, points with different admissibility structures have different distributions over $\mathcal{A}(x)$: they lead to different accessible futures. Therefore $\pi^{-1}(m)$ contains points with statistically distinguishable admissibility distributions.

Step 2. The projection π identifies all points in $\pi^{-1}(m)$. Points that are admissibility-distinct in X are representationally equivalent in M: they map to the same point m and are indistinguishable to any statistic computed in M.

Step 3. The admissibility partition $\{U_i\}$ assigns admissibility-distinct points to different classes. Under the identification produced by π, points from different admissibility classes are present in the same representational neighborhood U. The joint distribution $\pi_*\mu_U$ therefore contains contributions from multiple admissibility classes that the faithful distribution $\prod_i \pi_*\mu_{U_i}$ would keep separate.

Step 4. The mixing functional $\Lambda(m) > 0$ follows from the positivity of the information divergence D when the pushforward distribution fails to factorize. The lower bound by $f(\kappa_\pi(m))$ follows from continuity of Λ with respect to $\pi_*\mu$ and compactness of the fiber measure: as $\kappa_\pi(m)$ increases from zero, the within-fiber divergence between admissibility distributions increases continuously, and by continuity of Λ, so does the mixing functional. The function f is the composition of this two-step monotone relationship. $\square$

10.5 Statement: Strong Form

Theorem 10.5 (Projection-Collapse Principle, Strong Form)**.** *Under the additional regularity conditions:*

(i) Bounded fiber diameter*: $\sup_{x,x' \in \pi^{-1}(m)} d_g(x, x') \leq D < \infty$ for all $m \in M$;*

(ii) Lipschitz continuity*: $p(\cdot \mid x)$ is Lipschitz continuous in x with respect to the metric g, with constant L_p;*

(iii) Jensen-Shannon divergence*: the divergence D in* (9.1) *is the Jensen-Shannon divergence D_{JS};*

the mixing functional and the projected sectional curvature are asymptotically equivalent up to multiplicative constants:

$$\Lambda(m) \asymp \kappa_\pi(m). \tag{10.6}$$

Conjecture 10.6 (Strong Form Conjecture)**.** *The asymptotic equivalence* (10.6) *holds under the stated regularity conditions. The proof requires establishing a matching upper bound $\Lambda(m) \lesssim \kappa_\pi(m)$, which follows from the Lipschitz condition on p and the bounded fiber diameter via a Jensen-Shannon tensorization argument. The key steps are: the Lipschitz condition implies that within-fiber admissibility divergence grows at most linearly with fiber* *diameter; bounded fiber diameter caps the divergence; the Jensen-Shannon divergence is bounded and symmetric and satisfies a data-processing inequality that converts within-fiber divergence to representational mixing. This conjecture is presented as a research target for the companion analysis.*

10.6 The Regularity Conditions and Their Semantic Content

Each of the three regularity conditions in Theorem 10.5 has a precise semantic interpretation.

Bounded fiber diameter means that the ambiguity introduced by projection is bounded: the representation does not collapse infinitely many admissibility-distinct states into one point. This is a non-trivial requirement. It fails, for instance, when the encoding map π is highly lossy and many semantically incompatible states share a single representation. In practice, bounded fiber diameter corresponds to a requirement that the representational system has sufficient capacity to distinguish the main branches of the admissibility structure.

Lipschitz continuity of $p(\cdot \mid x)$ in x means that the admissibility structure varies without discontinuous jumps: there are no sharp semantic phase transitions where the accessible futures change catastrophically over a small change in position. Domains with genuine discontinuities—sharp boundaries between legal and illegal moves in formal games, for instance—violate this condition and require a separate analysis.

Jensen-Shannon divergence, in place of KL divergence, is required because the KL divergence can be infinite when the support of $\pi_* \mu_U$ does not contain the support of $\prod_i \pi_* \mu_{U_i}$. Jensen-Shannon divergence is bounded by $\log 2$, symmetric, and metrizes weak convergence. These properties are essential for the tensorization argument that gives the upper bound in Theorem 10.5.

10.7 Interpretive Consequences

Theorem 10.4 establishes that collapse is a geometric necessity under projection with hidden curvature: it is not an implementation failure, not a training artifact, and not a property that can be eliminated by adding more representational capacity without changing the projection geometry. Any map $\pi : X \to M$ that identifies points in a curved admissibility manifold will produce observable mixing in proportion to the curvature it destroys.

Theorem 10.5, if the conjecture is confirmed, establishes something stronger: collapse measurements are geometric observables. They encode curvature information up to multiplicative constants in the same sense that gravitational lensing encodes spacetime curvature. The mixing functional Λ is not merely bounded by κ_π. It is asymptotically equivalent to it. A complete field of Λ-measurements over M therefore determines κ_π up to a constant, and the inverse problem of recovering the admissibility geometry from collapse data becomes well-posed in the sense of geometric tomography.

Part IV

The Inverse Problem

Chapter 11

Collapse as Geometric Tomography

11.1 The Inverse Problem Formulated

The forward problem is now solved: given the admissibility manifold (X, g) and the projection $\pi : X \to M$, the Projection-Collapse Principle determines the mixing functional Λ over M. The inverse problem reverses this: given a representational manifold M, a field of mixing measurements Λ over M, and a projection π (known or partially known), recover the projected sectional curvature κ_π over the fibers of π, and from κ_π, the admissibility density $p(y \mid x)$ up to the resolution limit imposed by π.

This is an inverse problem in the sense of Tikhonov: the forward map $\kappa_\pi \mapsto \Lambda$ is known from the theorem; the task is to invert it from noisy, finite-resolution observations. Standard results from the theory of ill-posed inverse problems apply: uniqueness of the reconstruction requires additional structural assumptions, and stability of the reconstruction requires regularization.

11.2 Analogy with Geometric Tomography

The analogy with gravitational lensing and seismic tomography is precise enough to import mathematical structure from those domains. In gravitational lensing, the observable is the deflection of light passing through a gravitational field; the hidden quantity is the spacetime metric. In seismic tomography, the observable is the travel time of elastic waves through the earth; the hidden quantity is the elastic velocity field. In both cases, the forward map (from the hidden geometry to the observable) is known, and the inverse problem is to recover the geometry from measurements of the observable.

Here, the observable is the mixing functional Λ and the hidden quantity is the projected sectional curvature κ_π. The forward map is the Projection-Collapse Principle. The inverse problem has the same structure as lensing and seismology: a geometry is inferred from its effect on the observable trajectories that pass through it. The semantic analogue of light deflection is the distortion of representational neighborhoods under the admissibility geometry.

11.3 Resolution Limits and the Cramér-Rao Bound

The Fisher metric on X implies a fundamental resolution limit on any representation π. The Cramér-Rao inequality, applied to the statistical manifold (X, g), states that no estimator of the position $x \in X$ from observations in $M = \pi(X)$ can achieve variance smaller than the inverse of the Fisher information in the projected directions. This is a geometric lower bound on the resolution of π: the representation cannot distinguish admissibility structures more finely than the metric allows.

Below this resolution scale, the inverse problem is ill-posed: multiple distinct admissibility geometries produce the same pattern of collapse measurements. Regularization is required to select among the consistent solutions. The appropriate regularization encodes a prior over admissibility geometries, and the choice of prior reflects assumptions about the constraint structure of the domain. A maximum-entropy prior selects the flattest geometry consistent with the data. A smoothness prior selects the most slowly varying geometry. A constraint-consistent prior selects the geometry most compatible with known structural properties of the domain.

11.4 What Can and Cannot Be Recovered

A precise characterization of recoverable information from Λ measurements is as follows. The projected sectional curvature κ_π is recoverable up to fiber-averaging: the integral (8.2) can be estimated from Λ measurements, but the point-wise curvature $|K_g(v, h)|$ at individual points in the fiber cannot be resolved without additional measurements. The full Riemann tensor R_g is not recoverable from Λ alone: only the mixed components $K_g(v, h)$ for $v \in V_x$ and $h \in H_x$ contribute to leakage; the pure vertical and pure horizontal components require separate probes.

The admissibility density $p(y \mid x)$ is recoverable up to the equivalence class defined by the fiber structure of π: two density functions that agree on every fiber of π produce the same Λ field and cannot be distinguished by collapse measurements alone. Distinguishing within-fiber variation requires perturbation experiments—controlled changes to the input that probe specific tangent directions—or contrastive evaluations that compare the accessible futures of adjacent positions in the fiber.

11.5 The Reconstruction Operator

The inverse problem can be formalized as a bounded linear operator between function spaces. Let $L^2(M, \mu_M)$ denote the space of square-integrable functions on the representational manifold with respect to the pushforward measure $\mu_M = \pi_*\mu$. Define the *reconstruction operator*:

$$\mathcal{R} : L^2(M, \mu_M) \to L^2(M, \mu_M), \qquad \mathcal{R}(\Lambda) = \kappa_\pi, \tag{11.1}$$

as the operator that maps an observed mixing field Λ to the projected sectional curvature field κ_π that produced it through the Projection-Collapse Principle.

Theorem 11.1 (Stability of Reconstruction). *Under the regularity conditions of Theorem 10.5 (bounded fiber diameter D, L_p-Lipschitz admissibility density, Jensen-Shannon divergence), the reconstruction operator $\mathcal{R}$ satisfies the stability estimate:*

$$\|\mathcal{R}(\Lambda_1) - \mathcal{R}(\Lambda_2)\|_{L^2(M)} \le C_{D,L_p} \|\Lambda_1 - \Lambda_2\|_{L^2(M)}, \tag{11.2}$$

for a constant C_{D,L_p} depending on the fiber diameter bound D and the Lipschitz constant L_p. Consequently, small perturbations in the observed mixing field produce bounded perturbations in the reconstructed curvature.

Proof sketch. From the asymptotic equivalence $\Lambda(m) \asymp \kappa_\pi(m)$ of Theorem 10.5, the ratio $\kappa_\pi(m)/\Lambda(m)$ is bounded above and below by constants depending on D and L_p. Therefore:

$$\begin{aligned}\|\mathcal{R}(\Lambda_1) - \mathcal{R}(\Lambda_2)\|_{L^2}^2 &= \int_M |\kappa_\pi^{(1)}(m) - \kappa_\pi^{(2)}(m)|^2 \, d\mu_M(m) \\ &\le C^2 \int_M |\Lambda_1(m) - \Lambda_2(m)|^2 \, d\mu_M(m) = C^2 \|\Lambda_1 - \Lambda_2\|_{L^2}^2,\end{aligned}$$

where the inequality uses the pointwise bound $|\kappa_\pi^{(1)} - \kappa_\pi^{(2)}| \le C\,|\Lambda_1 - \Lambda_2|$ from the asymptotic equivalence. The constant $C = C_{D,L_p}$ is the one from Theorem 10.5. □

Theorem 11.1 establishes that the inverse problem is well-posed in the Hadamard sense: a solution exists (by the Projection-Collapse Principle), it is unique up to fiber equivalence (by the separating properties of D_{JS}), and it depends continuously on the data (by the stability estimate (11.2)). The reconstruction is therefore not merely an aspiration but a certifiably stable inverse operation.

The stability estimate also quantifies the resolution limit of the tomographic reconstruction: uncertainty in Λ by ϵ translates into uncertainty in κ_π by at most $C_{D,L_p}\,\epsilon$. To reconstruct curvature to precision δ, it suffices to measure the mixing functional to precision $\delta/C_{D,L_p}$.

Chapter 12

Spectral Reconstruction Theory

The reconstruction operator $\mathcal{R}$ of the previous chapter is shown to exist and be stable, but its construction is left implicit. This chapter makes the construction explicit by developing a spectral theory for the forward operator $L : \kappa_\pi \mapsto \Lambda$. The spectral decomposition transforms the reconstruction from an abstract inverse operator into a concrete computational procedure, and identifies the ill-posedness structure that regularization must address.

12.1 The Linearized Forward Operator

In the regime where Λ depends approximately linearly on κ_π—which holds in the asymptotic regime of Theorem 10.5 where $\Lambda \asymp \kappa_\pi$—define the *linearized forward operator*:

$$L : L^2(M, \mu_M) \to L^2(M, \mu_M), \qquad L\kappa_\pi = \Lambda. \tag{12.1}$$

The operator L encodes the geometric relationship between the hidden curvature field and the observable mixing field. Explicitly, L is an integral operator:

$$(L\kappa_\pi)(m) = \int_M K(m, m')\, \kappa_\pi(m')\, d\mu_M(m'), \tag{12.2}$$

where the kernel $K(m, m')$ describes how curvature at m' affects mixing at m through the fiber geometry. The kernel is determined by the projection π and the Fisher metric g: points in the same fiber contribute directly, while points in nearby fibers contribute through the A-tensor coupling (8.6).

Proposition 12.1 (Compactness and Self-Adjointness)**.** *If the admissibility manifold X is compact and the kernel $K(m, m')$ is square-integrable over $M \times M$, then:*

(i) L is compact as an operator on $L^2(M, \mu_M)$;

(ii) *L is self-adjoint when $K(m, m') = K(m', m)$, which holds when the fiber geometry is symmetric (the measure μ_m is symmetric under fiber automorphisms).*

Compactness of L is the precise statement of ill-posedness: the spectrum of L has zero as an accumulation point, meaning that the inverse operator L^{-1} is unbounded. Small errors in Λ can be amplified without bound by naive inversion.

12.2 Singular Value Decomposition

By the spectral theorem for compact self-adjoint operators, L has a countable orthonormal eigenbasis $\{\phi_n\}_{n\geq 1}$ of $L^2(M, \mu_M)$ with corresponding eigenvalues $\sigma_n \geq 0$ decreasing to zero:

$$L = \sum_{n=1}^{\infty} \sigma_n \, \phi_n \otimes \phi_n, \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq 0, \quad \sigma_n \to 0. \tag{12.3}$$

Expanding $\kappa_\pi = \sum_n a_n \phi_n$ and $\Lambda = \sum_n b_n \phi_n$ in this basis:

$$b_n = \sigma_n \, a_n, \qquad n \geq 1.$$

The naive inversion $a_n = b_n / \sigma_n$ is unstable when σ_n is small: small measurement errors in b_n produce large errors in a_n.

12.3 Curvature Recovery via Truncated SVD

The standard regularization strategy for compact operator inversion is truncated SVD:

Definition 12.2 (Truncated SVD Reconstruction)**.** For a truncation parameter $N \geq 1$, define the truncated reconstruction operator:

$$\mathcal{R}_N(\Lambda) = \sum_{n=1}^{N} \frac{\langle \Lambda, \phi_n \rangle_{L^2}}{\sigma_n} \, \phi_n. \tag{12.4}$$

Theorem 12.3 (Reconstruction Error for Truncated SVD)**.** *Let $\kappa_\pi = \sum_n a_n \phi_n$ be the true curvature field and $\tilde{\Lambda} = L\kappa_\pi + \epsilon_\Lambda$ the observed mixing field with measurement error*

$\|\epsilon_\Lambda\|_{L^2} \leq \epsilon$. *The truncated SVD reconstruction satisfies:*

$$\|\mathcal{R}_N(\tilde{\Lambda}) - \kappa_\pi\|_{L^2}^2 \leq \frac{\epsilon^2}{\sigma_N^2} + \sum_{n>N} a_n^2. \tag{12.5}$$

The first term is the noise amplification *(decreasing in N) and the second is the* truncation bias *(increasing in N). The optimal N balances these two terms.*

Proof. $\mathcal{R}_N(\tilde{\Lambda}) - \kappa_\pi = \sum_{n=1}^N \frac{\langle \epsilon_\Lambda, \phi_n \rangle}{\sigma_n} \phi_n - \sum_{n>N} a_n \phi_n$. The squared L^2 norm of the first sum is bounded by $\|\epsilon_\Lambda\|^2/\sigma_N^2$ (using $\sigma_n \geq \sigma_N$ for $n \leq N$); the second sum has squared norm $\sum_{n>N} a_n^2$. □

12.4 Tikhonov Regularization

An alternative to truncated SVD that avoids the hard threshold is Tikhonov regularization:

Definition 12.4 (Tikhonov Reconstruction)**.** For regularization parameter $\alpha > 0$, define:

$$\mathcal{R}_\alpha(\Lambda) = (L^*L + \alpha I)^{-1} L^* \Lambda, \tag{12.6}$$

where $L^* = L$ (self-adjoint case). In the eigenbasis:

$$\mathcal{R}_\alpha(\Lambda) = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2 + \alpha} \langle \Lambda, \phi_n \rangle \phi_n.$$

The Tikhonov filter $\sigma_n/(\sigma_n^2 + \alpha)$ smoothly suppresses small singular values: when $\sigma_n \gg \sqrt{\alpha}$, the filter is approximately $1/\sigma_n$ (faithful inversion); when $\sigma_n \ll \sqrt{\alpha}$, the filter is approximately σ_n/α (strong damping). The parameter α controls the effective truncation.

Theorem 12.5 (Tikhonov Error Bound)**.** *Under the source condition $\kappa_\pi = L\omega$ for some $\omega \in L^2(M, \mu_M)$ (which holds when κ_π is in the range of L), the Tikhonov reconstruction satisfies:*

$$\|\mathcal{R}_\alpha(\tilde{\Lambda}) - \kappa_\pi\|_{L^2} \leq \frac{\epsilon}{2\sqrt{\alpha}} + \frac{\alpha^{1/2}}{2}\|\omega\|_{L^2}, \tag{12.7}$$

minimized at the optimal $\alpha^* = (\epsilon/\|\omega\|_{L^2})^{2/3}$, giving error $O(\epsilon^{2/3})$.

The $O(\epsilon^{2/3})$ convergence rate is the classical Tikhonov rate for linear inverse problems with compact operators under source conditions. It is slower than the $O(\epsilon)$ rate for well-posed problems but is optimal for this class of ill-posed problems.

12.5 The Singular Value Spectrum as Semantic Resolution Limit

The singular value spectrum $\{\sigma_n\}$ of L encodes the resolution structure of the reconstruction. Large singular values correspond to coarse-scale curvature features that are reliably recovered from Λ measurements; small singular values correspond to fine-scale features that are suppressed by the projection and difficult to reconstruct.

The rate of decay of σ_n is determined by the geometry of the projection π and the smoothness of the admissibility density $p(\cdot \mid x)$. For a smooth submersion π with compact fibers, the singular values decay polynomially: $\sigma_n \asymp n^{-s/d}$ where s is the smoothness order of K and $d = \dim M$. For less regular projections, the decay can be faster (exponential for analytic K) or slower.

The semantic interpretation is direct: the resolution limit of the reconstruction is determined by the geometric structure of the projection. A projection that identifies many admissibility-distinct states (large fibers, high κ_π) has a more rapidly decaying singular value spectrum and is harder to invert. A projection that is nearly injective (small fibers, low κ_π) has slowly decaying singular values and admits accurate reconstruction. The spectral theory of L therefore quantifies, in a single compact form, the fundamental tradeoff between representational compression and admissibility reconstruction fidelity.

12.6 Triplet Measurements as Curvature Probes

A triplet (A, S, D) consists of an anchor state A, a valid substitute S, and a misleading decoy D. In a faithful representation, the substitute is more similar to the anchor than the decoy: $\text{sim}(A, S) > \text{sim}(A, D)$. A collapse event occurs when this inequality reverses: $\text{sim}(A, D) \geq \text{sim}(A, S)$.

In the geometric framework, a triplet collapse event corresponds to a location where the projected sectional curvature κ_π is large enough that the decoy D and the substitute S exchange representational proximity after projection. The curvature at the fiber over $\pi(A)$ is sufficient to rotate the accessibility profiles of S and D relative to A, placing D closer to A in M than S is, even though S is more similar to A in the admissibility geometry. Triplet evaluations are therefore point measurements of the sign and approximate magnitude of a curvature-induced proximity distortion.

The collapse rate $\mathrm{CR} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}[\mathrm{sim}(A_i, D_i) \geq \mathrm{sim}(A_i, S_i)]$ is an estimator of the proportion of locations in M where the projected curvature exceeds the proximity threshold. A battery of triplets, distributed across different operator classes and semantic domains, provides a coarse-grained curvature map of the admissibility manifold.

12.7 Entropy Measurements as Curvature Integrals

The entropy of a token's neighborhood, defined as $H(t) = -\sum_{u \in N_t} q(u \mid t) \log q(u \mid t)$ where N_t is the k-nearest-neighbor set and q is a softmax over neighbors by cosine similarity, is an integrated curvature measurement over a neighborhood of t in M. High entropy indicates that the neighborhood contains points from multiple admissibility classes—that the fiber over that region of M contains admissibility-distinct points. Low entropy indicates that the neighborhood is admissibility-coherent.

The typed entropy $H_C(t)$, obtained by restricting the sum to neighbors of type C, measures curvature within a single admissibility class. The cross-type leakage $\lambda(t) = \sum_{C \neq C_t} \sum_{u \in N_t \cap C} q(u \mid t)$ measures the total probability mass that has leaked across class boundaries—a direct estimate of Λ for a specific neighborhood. The drift index $\Delta H(t) = H(t) - \mathrm{median}_{s \in R(t)} H(s)$, where $R(t)$ is a role-matched reference set, measures the excess entropy relative to what would be expected from the admissibility structure alone.

These quantities form a hierarchy of increasingly coarse-grained curvature observables. Triplet measurements probe point values of proximity distortion. Entropy measurements probe integrated curvature over neighborhoods. Surface pressure $\pi(\Omega) = \mathbb{E}_{t \in \Omega}[\lambda(t)]$ measures average mixed curvature over an extended region. Together, they constitute the observational toolkit for the tomographic reconstruction of the admissibility manifold.

12.8 Perturbation Experiments as Tangent Vector Probes

Controlled perturbations of inputs—paraphrase, negation, modal substitution, tense shift, agent-patient reversal—probe the admissibility geometry in specific tangent directions. If a perturbation in direction $v \in T_x X$ produces large changes in the mixing functional Λ, then v has large components in the vertical bundle V_x and the projection is destroying information in that direction.

More precisely, a perturbation experiment measures the directional derivative of Λ with respect to the perturbation direction: $\partial_v \Lambda(\pi(x)) \approx \frac{\Lambda(\pi(x+\epsilon v))-\Lambda(\pi(x))}{\epsilon}$. Directions for which this derivative is large are the semantically sensitive directions: small displacements in those directions produce large changes in the admissibility structure. These are precisely the directions of high metric weight in g that the projection should preserve but may not.

Perturbation experiments are therefore measurements of the vertical-horizontal decomposition at specific points. They identify which tangent directions at x the representation is collapsing, and they provide gradient information that can be used to refine the tomographic reconstruction.

12.9 Combining Probes: Toward Tomographic Reconstruction

The three classes of measurement—triplet evaluations, entropy measurements, and perturbation experiments—provide complementary information about the admissibility geometry. Triplets are sparse but precise: they give point estimates of proximity distortion at specific locations. Entropy measurements are dense but coarse: they integrate curvature over neighborhoods and cannot resolve fine-grained structure. Perturbation experiments are directional: they probe specific tangent directions and provide gradient information unavailable from the other two.

A tomographic reconstruction combines all three. Triplet violations locate the curvature concentrations. Entropy gradients provide coarse-grained curvature maps. Perturbation derivatives identify the directions in which the curvature is concentrated. Together, under sufficient density of measurements and appropriate regularization, they constrain the admissibility geometry up to the resolution limit imposed by the Cramér-Rao bound.

Chapter 13

Connection to Reentrant Value Field Theory

13.1 Bohlen's Framework and the Admissibility Manifold

Recent work on retarded functional differential equations (RFDEs) as a foundation for synthetic cognition provides a concrete instantiation of the abstract admissibility geometry developed in Parts II and III [6]. Bohlen constructs a coupled dynamical system in which a symbolic field H_L and a geometric field X_R, each a section of a vertex bundle over a finite graph, are coupled through a bipartite Hilbert-Schmidt operator with propagation delays. The central formal results—well-posedness, compact global attractor, and delay-independent stability—are proven under explicit admissibility conditions specified through nine synthetic design blueprints.

The connection to the present framework runs through several specific correspondences. The history space $\mathcal{X} = C([-\tau_{\max}, 0], \mathcal{Z})$ is Bohlen's version of the admissibility manifold X: it is the space of all admissible histories of the system, equipped with the sup-norm topology. The projection from the history space to the current state space $\pi : \mathcal{X} \to \mathcal{Z}$ given by evaluation at $t = 0$ is an instance of the general projection $\pi : X \to M$. Two history segments that agree at $t = 0$ but differ over $[-\tau_{\max}, 0)$ project to the same current state but lead to different future dynamics—precisely the structure of the fiber $\pi^{-1}(z)$.

13.2 The Small-Gain Condition as a Collapse Surface

The stability condition $C_K^2 < \mu_L \mu_R$ in Bohlen's framework has a natural interpretation in the admissibility geometry. Here C_K is the Hilbert-Schmidt norm of the interfield coupling operator and μ_L, μ_R are the dissipativity constants of the symbolic and geometric fields respectively. The condition is the positivity requirement for the 2×2 Schur margin matrix:

$$\begin{pmatrix} \mu_L & -C_K \\ -C_K & \mu_R \end{pmatrix} > 0 \iff C_K^2 < \mu_L \mu_R.$$

When this condition holds, the coupled system is globally asymptotically stable for all finite delays. When it is violated, the coupled system loses stability.

In the language of admissibility geometry, the stability boundary $C_K^2 = \mu_L \mu_R$ is a collapse surface in the state space: a region where the admissibility geometry becomes critically curved. Points in the history space that are near the stability boundary have large projected sectional curvature κ_π: small perturbations of the coupling strength produce large changes in the accessible futures of the system. The small-gain condition is therefore a condition on the admissibility geometry of the history-space projection, not merely a stability criterion in the conventional dynamical-systems sense.

13.3 The Lyapunov-Krasovskii Functional as Admissibility Density

The Lyapunov-Krasovskii functional:

$$V(\tilde{\phi}) = \frac{1}{2}\|\tilde{\phi}_L(0)\|_F^2 + \frac{1}{2}\|\tilde{\phi}_R(0)\|^2 + \frac{C_K^2}{2\mu_L}\int_{-\tau_{R\to L}}^{0} \|\tilde{\phi}_R(s)\|^2\,ds + \frac{C_K^2}{2\mu_R}\int_{-\tau_{L\to R}}^{0} \|\tilde{\phi}_L(s)\|_F^2\,ds \tag{13.1}$$

is a candidate for the admissibility density $\Phi(x)$ in the neighborhood of the equilibrium Z^*. It is a quadratic form on the history space that decreases along solutions—a Lyapunov function in the classical sense. The Hessian of V near the equilibrium is well-defined and positive definite precisely because V is constructed to be convex in that region.

This is one of the locations where the Hessian construction does work, but only locally. Near the equilibrium, where V is convex, the Hessian of V approximates the Fisher metric of the admissibility density $p(y \mid x)$ restricted to the attractor. The Fisher metric and the Hessian metric agree at the equilibrium and diverge away from it, with the Hessian failing at saddle points and the Fisher metric remaining well-defined throughout. Bohlen's framework provides a local approximation to the admissibility geometry near the attractor; the Fisher metric extends this approximation globally.

13.4 Blueprint Admissibility as Constraint Geometry

The nine synthetic design blueprints in Bohlen's framework are admissibility contracts: explicit conditions that each architectural component must satisfy for the system to remain within the admissible region of the state space. Each blueprint specifies regularity, boundedness, dissipativity, and symmetry conditions that, collectively, ensure that the master RFDE satisfies the hypotheses of the formal theorems.

In the language of the present framework, the blueprint conditions are sufficient conditions for the admissibility geometry to be non-degenerate. Blueprint 1 (symbolic update) ensures that the symbolic field evolves within a compact, positively invariant domain with a spectral gap—ensuring that the Fisher metric on the symbolic component of the history space has bounded curvature. Blueprint 3 (interconnector) bounds the Hilbert-Schmidt norm C_K of the coupling kernel—bounding the mixed curvature between the symbolic and geometric components of the history space. A violation of any blueprint condition corresponds to a region of high κ_π in the admissibility geometry: the system leaves the region where the constraint structure supports coherent future trajectories.

Chapter 14

Semantic Field Theory

14.1 From Diagnostics to Dynamics

The inverse problem formulated in Chapter 8 recovers a static picture of the admissibility geometry: given Λ measurements over M, reconstruct κ_π over the fibers of π. But semantic systems are dynamic. They evolve under optimization, learning, and use. A complete theory requires dynamics: equations of motion for the admissibility field under these pressures.

The central observation is that compression-driven optimization acts as a force that flattens the admissibility field. Any optimizer that reduces representational complexity by minimizing the distance between similar representations will, unless specifically constrained, collapse distinctions that are not locally useful for the current task. This is not a failure mode. It is the expected behavior of an optimizer that acts on the representation rather than on the underlying constraint field. Collapse is what optimization looks like from the perspective of the admissibility geometry.

14.2 RSVP as a Field-Theoretic Framework

The Relativistic Scalar-Vector Plenum (RSVP) framework provides a candidate dynamics for the admissibility field. In RSVP, the primary objects are a scalar field Φ and a vector field $\mathbf{v}$ over a spacetime manifold, coupled through field equations that enforce the constraint structure of the domain. The connection to the Fisher geometry developed in this monograph can be made precise through the partition function construction.

From Admissibility Density to Gibbs Distribution

Define the scalar admissibility potential:

$$\Phi(x) = \log |\mathcal{A}(x)|, \tag{14.1}$$

the log-volume of the admissible future cone at x. Given an energy function $E(y, x) \geq 0$ measuring the cost of continuation y from position x, define the Gibbs-type admissibility density:

$$p(y \mid x) = \frac{\exp(-E(y,x))}{Z(x)}, \qquad Z(x) = \int_{\mathcal{A}(x)} \exp(-E(y,x))\, dy, \tag{14.2}$$

where $Z(x)$ is the partition function. The RSVP scalar field is identified with the log-partition function:

$$\Phi(x) = \log Z(x). \tag{14.3}$$

Proposition 14.1 (Fisher Metric from the Partition Function)**.** *For the Gibbs density* (14.2)*, the Fisher information metric on the admissibility manifold X is:*

$$g_{ij}(x) = \partial_i \partial_j \log Z(x) - \frac{\partial_i Z(x)\, \partial_j Z(x)}{Z(x)^2}. \tag{14.4}$$

In the case where the energy $E(y, x)$ is linear in x—that is, $E(y, x) = -\langle \theta(x), \phi(y) \rangle$ for sufficient statistics $\phi(y)$ and natural parameters $\theta(x)$—the admissibility density is an exponential family and the metric simplifies to:

$$g_{ij}(x) = \partial_i \partial_j \log Z(\theta(x)), \tag{14.5}$$

the Hessian of the log-partition function in the natural parameters.

Proof. By direct computation from the Fisher metric definition (7.1):

$$\begin{aligned}
g_{ij}(x) &= \mathbb{E}_{y \sim p(\cdot|x)} \left[\partial_i \log p(y \mid x) \cdot \partial_j \log p(y \mid x) \right] \\
&= \mathbb{E}_{y \sim p(\cdot|x)} \left[(\partial_i(-E(y,x)) - \partial_i \log Z(x)) \left(\partial_j(-E(y,x)) - \partial_j \log Z(x) \right) \right] \\
&= \mathrm{Cov}_{y \sim p(\cdot|x)} \left[\partial_i(-E(y,x)),\, \partial_j(-E(y,x)) \right],
\end{aligned}$$

which equals the covariance of the score function, the standard information-geometric expression (14.4). For the exponential family case $\log p(y \mid x) = \langle \theta(x), \phi(y) \rangle -$

$\log Z(\theta(x))$, the score function is $\partial_i \log p = \partial_i \theta^k \cdot \phi_k - \partial_i \log Z$, and the Fisher metric in the natural parameterization is the Hessian of $\log Z$, which is positive definite wherever Z is strictly log-convex. $\square$

Proposition 14.1 establishes the precise bridge between RSVP field theory and the Fisher geometry. The RSVP scalar field $\Phi = \log Z$ is not merely an analogy for admissibility density. It is the generating function of the Fisher metric: its Hessian in natural coordinates is the metric tensor. The RSVP field equations, which govern the dynamics of Φ, therefore govern the dynamics of the Fisher metric and, through the Projection-Collapse Principle, the dynamics of collapse.

Collapse as RSVP Field Deformation

In the RSVP framework, the field equations:

$$\Box\Phi = \lambda(\Phi^2 - v^2)\Phi - \nabla_\mu V^\mu, \tag{14.6}$$

$$\nabla_\nu F^{\mu\nu} = J^\mu[\Phi], \tag{14.7}$$

enforce a Mexican-hat potential for Φ, producing spontaneous symmetry breaking at locations where the admissibility density $Z(x)$ falls below the vacuum expectation v. These are the semantic bottlenecks of Chapter 3: locations where the admissible future cone contracts sharply.

Optimization that violates these equations deforms Φ away from its equilibrium profile, flattening the partition function landscape and destroying the curvature that the Fisher metric encodes. Compression-driven learning—gradient descent on a loss that rewards representational compactness—is exactly such a deformation: it drives $Z(x)$ toward uniformity, reducing the variance of $p(\cdot \mid x)$ across fibers and suppressing the mixed sectional curvature κ_π. The Projection-Collapse Principle describes the observable consequence: the mixing functional Λ decreases as κ_π is flattened, and the reconstruction program becomes correspondingly less informative.

This analysis suggests a dynamical interpretation of the monograph's program: the reconstruction operator $\mathcal{R}$ of Theorem 11.1 can be applied not only to a static field of measurements but to a time series of measurements, tracking the deformation of the admissibility geometry over the course of training. Collapse curves—plots of Λ against training step—become trajectories in the space of RSVP field configurations, and the reconstruction program recovers the underlying field-theoretic dynamics from the observable collapse trace.

14.3 Meaning as a Field-Theoretic Invariant

The conclusion of the theoretical development can now be stated precisely. Meaning is not a property of representations. It is an invariant of the admissibility field: a quantity preserved under admissibility-respecting transformations and destroyed by transformations that violate the constraint structure.

More precisely: a semantic transformation $T : X \to X$ is admissibility-respecting if it maps admissible continuations to admissible continuations: $T(\mathcal{A}(x)) \subseteq \mathcal{A}(T(x))$ for all $x \in X$. Under admissibility-respecting transformations, the Fisher metric is preserved: $T^* g = g$. The Riemann curvature tensor is preserved: $T^* R_g = R_g$. The Projection-Collapse Principle is preserved: the relationship between Λ and κ_π holds in the transformed coordinates. Semantic integrity is the preservation of these geometric invariants.

Semantic collapse is the observable signature of a transformation that is not admissibility-respecting: one that maps some admissible continuations outside the admissible region, flattening the curvature of the admissibility manifold and producing observable mixing in the representational space. The collapse measurements Λ are the projected residue of this geometric violation.

Chapter 15

Variational RSVP Geometry

The RSVP scalar field $\Phi = \log Z$ generates the Fisher metric in the exponential family setting (Proposition 14.1). This chapter develops the full variational formulation: the admissibility action whose Euler-Lagrange equations are the RSVP field equations, the curvature evolution under these dynamics, and the interpretation of symmetry breaking as collapse formation.

15.1 The Admissibility Action

Define the *admissibility action functional* on the space of scalar fields $\Phi : X \to \mathbb{R}$:

$$S[\Phi] = \int_X \left(\frac{1}{2} |\nabla \Phi|_g^2 - V(\Phi) \right) d\mu_g, \tag{15.1}$$

where $|\nabla \Phi|_g^2 = g^{ij} \partial_i \Phi \, \partial_j \Phi$ is the squared norm of the gradient in the Fisher metric, $V : \mathbb{R} \to \mathbb{R}$ is a potential function, and $d\mu_g$ is the Riemannian volume form of (X, g).

The action (15.1) has two terms. The kinetic term $\frac{1}{2} |\nabla \Phi|_g^2$ penalizes rapid variation of the admissibility density across the manifold: regions where Φ changes quickly contribute large kinetic energy, and the action favors smooth admissibility profiles. The potential term $V(\Phi)$ encodes the preferred admissibility level: the specific shape of V determines the equilibrium distribution of Φ.

15.2 Euler-Lagrange Equations: The Admissibility Field Equation

Theorem 15.1 (Admissibility Field Equation)**.** *Critical points of the action $S[\Phi]$ satisfy the Euler-Lagrange equation:*

$$\Delta_g \Phi = V'(\Phi), \tag{15.2}$$

where $\Delta_g = g^{ij}\nabla_i\nabla_j$ is the Laplace-Beltrami operator on (X, g).

Proof. Consider a variation $\Phi_\epsilon = \Phi + \epsilon\eta$ for compactly supported $\eta : X \to \mathbb{R}$. Then:

$$\frac{d}{d\epsilon}\bigg|_0 S[\Phi_\epsilon] = \int_X \left(g^{ij}\partial_i\Phi\,\partial_j\eta - V'(\Phi)\eta \right) d\mu_g.$$

Integrating the first term by parts (using the divergence theorem on (X, g) with compact support of η):

$$\int_X g^{ij}\partial_i\Phi\,\partial_j\eta\, d\mu_g = -\int_X \eta\,\Delta_g\Phi\, d\mu_g.$$

Setting the variation to zero for all η gives (15.2). □

The field equation (15.2) is the admissibility analogue of the Laplace or Poisson equation. It determines the equilibrium admissibility profile: at equilibrium, the Laplacian of Φ (measuring how Φ deviates from its local average) equals the derivative of the potential (the force pulling Φ toward the potential minimum).

15.3 The Mexican-Hat Potential and Semantic Phase Transitions

The RSVP framework uses a Mexican-hat (double-well) potential:

$$V(\Phi) = \frac{\lambda}{4}(\Phi^2 - v^2)^2, \tag{15.3}$$

with $\lambda > 0$ and $v > 0$ the vacuum expectation value. This potential has two minima at $\Phi = \pm v$ and a local maximum at $\Phi = 0$.

Proposition 15.2 (Symmetry Breaking and Bottleneck Formation). *For the potential* (15.3)*, the field equation* (15.2) *becomes:*

$$\Delta_g \Phi = \lambda \Phi(\Phi^2 - v^2). \tag{15.4}$$

Regions where $|\Phi| < v$ are unstable under the dynamics: the potential force $\lambda\Phi(\Phi^2 - v^2)$ drives Φ toward $\pm v$. Regions where $\Phi \to 0$ are semantic bottlenecks: the admissibility density collapses toward zero, and the accessible future cone contracts to a point.

This identifies semantic phase transitions geometrically: when the admissibility field Φ crosses the unstable region $|\Phi| < v$, the system undergoes a transition from a semantically rich region (many accessible futures, $\Phi \approx v$) to a bottleneck ($\Phi \approx 0$). The transition corresponds to the formation of a collapse surface in the sense of Chapter 5: a region of high projected sectional curvature that generates large observable mixing.

15.4 Fisher-RSVP Correspondence: The Metric from the Action

The connection between the action (15.1) and the Fisher metric is made precise through the following correspondence.

Theorem 15.3 (Fisher-RSVP Correspondence). *For an exponential family admissibility density $p(y \mid x) = \exp(\langle \theta(x), \phi(y) \rangle - \Phi(\theta(x)))$ with natural parameters $\theta(x)$ and log-partition function $\Phi = \log Z$, the Fisher metric is:*

$$g_{ij}(x) = \partial_i \partial_j \Phi(\theta(x)) = \partial_i \partial_j \log Z(\theta(x)), \tag{15.5}$$

and the Riemann curvature tensor of g satisfies:

$$R_{ijkl} = \frac{1}{2}\left(\partial_i \partial_l g_{jk} + \partial_j \partial_k g_{il} - \partial_i \partial_k g_{jl} - \partial_j \partial_l g_{ik}\right) + g^{pq}(\Gamma_{il,p}\Gamma_{jk,q} - \Gamma_{ik,p}\Gamma_{jl,q}), \tag{15.6}$$

where all Christoffel symbols are computed from Φ via $g_{ij} = \partial_i \partial_j \Phi$.

Proof. The first statement is Proposition 14.1. The second statement is the standard formula for the Riemann tensor in terms of the metric and its derivatives, applied to $g_{ij} = \partial_i \partial_j \Phi$. □

Theorem 15.3 completes the RSVP-geometry bridge: the RSVP scalar field Φ is the generating function of the Fisher metric. The action $S[\Phi]$ is therefore an action on the space of Riemannian metrics on X (within the exponential family class), and the field equation $\Delta_g \Phi = V'(\Phi)$ is an equation of motion for the metric tensor itself. The RSVP field theory is, in the exponential family regime, a field theory on the space of admissibility geometries.

15.5 Curvature Evolution Under RSVP Dynamics

The field equation (15.2) determines how the metric $g = \partial^2\Phi$ evolves when Φ is perturbed. Differentiating (15.5) twice more:

$$\frac{\partial}{\partial t} g_{ij} = \partial_i \partial_j \dot{\Phi},$$

where $\dot{\Phi}$ satisfies the linearized field equation $\Delta_g \dot{\Phi} = V''(\Phi)\dot{\Phi}$. This is a Schrödinger-type equation on the admissibility manifold with potential $V''(\Phi)$. Near the vacuum $\Phi = \pm v$, the potential is $V''(\pm v) = 2\lambda v^2 > 0$, giving a massive scalar field: small perturbations of Φ are stable and decay exponentially. Near the bottleneck $\Phi = 0$, the potential is $V''(0) = -\lambda v^2 < 0$, giving a tachyonic scalar field: perturbations grow exponentially, corresponding to the onset of phase transition.

The curvature of g evolves according to the linearization of (15.6) in $\dot{\Phi}$. In particular, the scalar curvature $R = g^{ij} R_{ij}$ satisfies an evolution equation that couples to the Laplacian of $V''(\Phi)$. The collapse-driving curvature κ_π is a projection of R_{ijkl} and therefore inherits this dynamics. This gives a precise sense in which the RSVP field theory predicts the temporal evolution of collapse: as Φ evolves under (15.2), the mixing functional $\Lambda(t) \asymp \kappa_\pi(t)$ evolves according to the projected curvature dynamics.

Chapter 16

Persistent Admissibility Topology

The geometric theory developed in Parts II and III describes the local structure of the admissibility manifold: curvature at points, volume of balls, geodesic distances. But collapse and semantic integrity have global aspects that local geometry cannot capture. Disconnected admissibility basins, topological loops that create contextual memory effects, and voids in the accessible future space are topological features that persist across scales. This chapter applies persistent homology to the admissibility filtration to extract these global features.

16.1 The Admissibility Filtration

Definition 16.1 (Admissibility Filtration)**.** For the scalar admissibility field $\Phi : X \to \mathbb{R}$, define the *superlevel set filtration*:

$$X_\lambda = \{x \in X : \Phi(x) \geq \lambda\}, \qquad \lambda \in \mathbb{R}. \tag{16.1}$$

As λ decreases from $\sup \Phi$ to $-\infty$, the sets X_λ grow monotonically: $X_{\lambda'} \subseteq X_\lambda$ for $\lambda' \geq \lambda$. The collection $\{X_\lambda\}_{\lambda \in \mathbb{R}}$ is the admissibility filtration.

The filtration (16.1) tracks the topology of the admissibility manifold at each threshold λ. At high λ, only the most admissible positions are included: X_λ consists of isolated peaks of the admissibility landscape. As λ decreases, basins of admissibility merge, tunnels appear, and voids form. The homology groups $H_k(X_\lambda)$ track these topological changes.

16.2 Homology of the Admissibility Filtration

For each threshold λ, the homology groups $H_k(X_\lambda; \mathbb{Z})$ with $k = 0, 1, 2, \ldots$ capture topological features of X_λ:

$$\begin{aligned}
H_0(X_\lambda) &\cong \mathbb{Z}^{c(\lambda)}, \quad \text{where } c(\lambda) = \text{number of connected components of } X_\lambda; \\
H_1(X_\lambda) &\cong \mathbb{Z}^{\ell(\lambda)}, \quad \text{where } \ell(\lambda) = \text{number of independent loops in } X_\lambda; \\
H_2(X_\lambda) &\cong \mathbb{Z}^{v(\lambda)}, \quad \text{where } v(\lambda) = \text{number of independent voids in } X_\lambda.
\end{aligned}$$

The semantic interpretations are: Connected components of X_λ are disconnected admissibility basins: regions of semantic space with sufficient admissibility that are topologically isolated from each other at threshold λ. A system in one basin cannot reach another by a path that stays above the threshold.

Loops in X_λ are cyclic admissibility structures: paths that return to the same semantic position while remaining above threshold. These correspond to contextual memory effects—the holonomy discussed in the open problems chapter—where returning to the same representation along different admissibility routes yields different accessible futures.

Voids in X_λ are regions of admissibility space completely enclosed by the superlevel set but interior to it: holes in the accessible future. These are the locations where the constraint structure forbids certain continuations from all surrounding positions.

16.3 Persistence Diagrams

Definition 16.2 (Persistence Diagram)**.** A topological feature (connected component, loop, or void) is *born* at threshold λ_b and *dies* at threshold $\lambda_d < \lambda_b$ (the feature appears as λ decreases and disappears as λ decreases further). The *persistence* of the feature is $\lambda_b - \lambda_d$. The *persistence diagram* Dgm_k is the multiset of birth-death pairs (λ_b, λ_d) for all k-dimensional features.

Features with large persistence $\lambda_b - \lambda_d$ are geometrically significant: they survive a wide range of threshold values and correspond to robust structures in the admissibility landscape. Features with small persistence are geometrically noise: they appear and disappear within a narrow range and correspond to local fluctuations.

Theorem 16.3 (Stability of Persistence Diagrams [9]). *If $\Phi, \Psi : X \to \mathbb{R}$ are two admissibility fields with $\|\Phi - \Psi\|_\infty \leq \epsilon$, then the bottleneck distance between their persistence diagrams satisfies:*

$$d_B(\mathrm{Dgm}_k(\Phi), \mathrm{Dgm}_k(\Psi)) \leq \epsilon, \tag{16.2}$$

for all k. Small perturbations of the admissibility field produce small perturbations in the persistence diagram.

Theorem 16.3 means that the topological features of the admissibility filtration are robust: they can be estimated from approximations to Φ without catastrophic topological error, as long as the approximation error is small in the sup-norm.

16.4 Topological Bottlenecks and Disconnected Basins

A topological bottleneck is a birth-death pair (λ_b, λ_d) in Dgm_0 (the H_0 diagram): a connected component of X_{λ_b} that merges with another component at λ_d. The persistence $\lambda_b - \lambda_d$ measures the depth of the admissibility saddle between the two basins: to pass from one basin to the other while staying in X_λ, one must descend at least $\lambda_b - \lambda_d$ below the peak of the lower basin.

In semantic terms, a topological bottleneck of high persistence corresponds to two admissibility basins that are strongly separated: semantic positions in one basin have accessible futures that are radically different from those in the other, and moving between them requires passing through a region of very low admissibility. This is the global analogue of the local bottleneck of Definition 6.5: local bottlenecks correspond to high curvature, while topological bottlenecks correspond to persistent H_0 features.

16.5 Semantic Topological Invariants

The persistence diagrams $\{\mathrm{Dgm}_k\}$ are topological invariants of the admissibility field Φ. They are preserved under admissibility-respecting transformations (which preserve Φ) and changed by transformations that deform Φ. Semantic collapse, as a deformation of the admissibility field under compression-driven optimization, therefore changes the persistence diagrams: components merge, loops disappear, and voids fill in as Φ is flattened.

The persistence diagrams provide a global audit of semantic integrity that complements the local collapse diagnostics of Chapter 9. Where the collapse graph (Chapter 12) identifies local curvature concentrations, the persistence diagram identifies global topological structure: which admissibility basins are robustly separated, which loops create contextual memory effects, and which voids represent globally forbidden continuations.

The relationship between the persistence diagram and the reconstruction operator $\mathcal{R}$ of Theorem 11.1 is an open problem. In the geometric tomography analogy, persistence features of the seismic velocity field can be recovered from travel-time data; here, the question is whether H_0 persistence features of Φ can be recovered from Λ measurements. Theorem 16.3 suggests that they can be approximately recovered whenever $\mathcal{R}(\Lambda)$ approximates κ_π closely enough to determine Φ up to the sup-norm tolerance implied by the desired persistence resolution.

Part V

Architecture and Infrastructure

Chapter 17

Constraint-First Semantic Architectures

17.1 The Design Inversion

Standard architectures for semantic systems generate representations first and apply constraints afterward. The generation step produces a distribution over continuations, typically by predicting next tokens from a compressed context. The constraint step filters or reweights this distribution to enforce various properties: fluency, coherence, factuality, safety. The constraint step is downstream of generation. It is a repair mechanism.

The admissibility framework inverts this order. Admissibility is the primary computational object. Generation occurs by sampling from $p(\cdot \mid x)$, the conditional density over admissible continuations from the current semantic position $x \in X$. The distribution p encodes the constraint structure from the beginning. There is no separate repair step because inadmissible trajectories are never entered into the generative distribution. The constraint is the prior.

This inversion has consequences at every level of the architecture. The training objective changes: instead of minimizing prediction error on a fixed dataset, the system learns to approximate the admissibility density $p(y \mid x)$ of the constraint structure. The inference procedure changes: instead of sampling from a learned next-token distribution and filtering, the system samples from the approximated $p(\cdot \mid x)$ directly. The evaluation protocol changes: instead of measuring accuracy against a test set, the system is evaluated by the fidelity of its p-approximation, measured through triplet evaluations, entropy diagnostics, and perturbation experiments.

17.2 Admissibility Filtering as a Primitive Operation

In a constraint-first architecture, the fundamental operation is admissibility checking: given a proposed continuation y and a current position $x \in X$, is $y \in \mathcal{A}(x)$? This check is not a post-generation filter. It is the inner loop of the generative process. All computation is organized around maintaining the current position within the admissible region of X.

The difference between this and standard filtering is not superficial. A standard filter acts on the output of a generative process that is indifferent to admissibility. It discards inadmissible outputs after they have been generated, at a cost proportional to the probability mass placed on inadmissible regions. A constraint-first architecture never generates inadmissible outputs, because the generative distribution has zero support outside $\mathcal{A}(x)$. The gradient signal during training is concentrated on the admissible region. The failure modes are different: instead of generating fluent but inadmissible content, the system risks under-representing the boundary regions of $\mathcal{A}(x)$ where admissibility is uncertain.

17.3 Typed Representations and Projection Discipline

If representations are projections of admissibility geometry, they should carry information about their position in the vertical-horizontal decomposition. A typed representation assigns to each point $m \in M$ a label that identifies the admissibility class C to which it belongs and specifies which directions in $T_m M$ are horizontal (reliable) and which are unreliable because they are the projections of curved fiber directions.

Projection discipline is the requirement that inference depends only on horizontal features: features in directions that the projection faithfully transmits from the admissibility geometry. Inference that depends on vertical features is inference that depends on information the representation cannot reliably encode, because the vertical directions are subject to the curvature-induced distortions quantified by κ_π. Projection discipline is therefore a structural constraint on the architecture, not a post-hoc regularizer.

In Bohlen's framework, the analogue of projection discipline is the blueprint admissibility system: each component of the architecture is constrained to operate within a compact, positively invariant domain with specified dissipativity and symmetry properties. The blueprint conditions collectively ensure that the architecture remains in the region of the history space where the admissibility geometry is non-degenerate, and they specify which coupling directions are reliable (those satisfying the small-gain condition) and which are not.

Chapter 18

Diagnostics, Auditing, and Certification

18.1 The Diagnostic Hierarchy

A complete diagnostic system for admissibility geometry operates at multiple scales. Point measurements, provided by triplet evaluations, locate specific collapse events: they identify pairs of states (A, D) that have exchanged relative proximity to an anchor A under the admissibility geometry. Neighborhood measurements, provided by entropy and leakage statistics, integrate curvature over local regions: they identify areas of the representational space where the admissibility structure is globally unstable. Regional measurements, provided by surface pressure and collapse graphs, identify extended regions of high mixed curvature: the collapse surfaces where multiple admissibility classes simultaneously lose separation. Global measurements, provided by total curvature integrals and topological invariants of the admissibility manifold, characterize the overall geometry.

Each scale provides different information and is appropriate for different evaluation tasks. Point measurements are appropriate for targeted audits of specific semantic distinctions. Neighborhood measurements are appropriate for detecting general instability in a region of the representational space. Regional measurements are appropriate for identifying the structural weaknesses of an architecture—the locations where the projection geometry systematically fails. Global measurements are appropriate for comparing architectures or for tracking the evolution of admissibility geometry over the course of training.

18.2 Collapse Graphs as Audit Artifacts

Collapse graphs—distributions of triplet violations by operator class and region—are reinterpreted in the admissibility framework as curvature maps of the projected admissibility geometry. A collapse graph plots, for each location in M and each admissibility class, the empirical estimate of $\kappa_\pi(m)$ derived from triplet measurements. Regions of high collapse rate correspond to regions of high projected sectional curvature. The class-specific structure of the graph identifies which types of semantic distinction are most severely affected by the projection.

As audit artifacts, collapse graphs provide task-grounded, falsifiable evidence about where a system's representational compression is destroying admissibility structure. They are not summaries of average performance. They are maps of specific geometric failures. A system with a high overall accuracy but a collapse graph concentrated in the modal operator region is failing to maintain the modal distinctions on which logical inference depends, regardless of what the accuracy metric reports.

18.3 Certification via Proof Obligations

The admissibility framework supports formal certification of constraint invariants. The procedure mirrors the blueprint admissibility system of Bohlen's framework: export the operator graph G_C encoding the local admissibility structure, verify that it satisfies target frame properties (reflexivity, transitivity, Euclidean closure corresponding to the modal systems T, S4, S5), and use the verification as a certificate attached to the model.

In practice, this can be implemented using theorem provers (Isabelle, Coq) for small graphs exported from held-out checkpoints, or using SMT solvers (Z3) for fast screening of frame properties encoded as first-order conditions over adjacency matrices. Counterexamples produced by these solvers identify specific locations where the admissibility geometry violates the target invariants. These counterexamples are hard negatives for continued training: they identify the points in the representational space where the projection is destroying the most curvature, and they provide gradient signal for repairing the projection geometry.

Part VI

Foundations and Consequences

Chapter 19

Accessibility Fields, Trajectories, and Semantic Action

Before situating the framework within the broader literature, we develop three complementary mathematical structures that deepen the kinematic and dynamic content of the theory: the accessibility field as a primary object, the replacement of states by trajectory equivalence classes, and a semantic action functional that unifies local geometry, entropic volume, and constraint dynamics into a single variational principle.

19.1 Accessibility Fields

The admissibility manifold X carries more structure than a smooth manifold with a Riemannian metric. Associated to each point $x \in X$ is a set of admissible continuations $\mathcal{A}(x) \subseteq X$. The assignment:

$$\mathcal{A} : X \to \mathcal{P}(X), \qquad x \mapsto \mathcal{A}(x), \tag{19.1}$$

defines the *accessibility field*: a map from semantic positions to subsets of the manifold specifying which positions are reachable in a single admissible step. The probability density $p(\cdot \mid x)$ over $\mathcal{A}(x)$ introduced in Chapter 3 is a measurable refinement of this field, encoding not merely reachability but admissibility weight.

The accessibility field is the primary semantic object in the framework. The local geometry of meaning at x is determined not by the position of x in isolation but by the structure of $\mathcal{A}(x)$: what futures are open, how densely they are accessible, and how that structure changes as x varies. Two positions $x, x' \in X$ with identical coordinates but different accessibility fields have different meanings. Two positions with different coordinates but identical accessibility fields are semantically equivalent.

The Fisher metric g is induced by $\mathcal{A}$: it measures how rapidly the accessibility field changes as x moves. Large $g_{ij}(x)$ in direction i means that small displacements in that direction substantially reshape $\mathcal{A}(x)$. Metric degeneracy means the accessibility field is insensitive to motion in that direction. The admissibility manifold with its Fisher metric is therefore a metrized accessibility field, not merely a Riemannian manifold equipped with a semantic interpretation.

19.2 Accessibility Entropy

The scalar admissibility field $\Phi(x) = \log Z(x)$ introduced in the RSVP chapters encodes the log-volume of the accessible continuation space. We define the accessibility entropy directly from $\mathcal{A}(x)$:

$$S(x) = \log \mu(\mathcal{A}(x)), \tag{19.2}$$

where μ is the volume measure on X. This quantity measures how much future freedom the semantic position x admits: large $S(x)$ indicates many accessible continuations and high semantic flexibility; small $S(x)$ indicates a bottleneck where the constraint structure sharply restricts forward motion.

The accessibility entropy connects three parts of the framework. In the kinematic theory, $S(x)$ is small precisely at the bottleneck points of Definition 6.5: where $V(x,r) \ll V_0(r)$, the Fisher ball is small, and the log-volume $S(x)$ is correspondingly suppressed. In the dynamic theory, $S(x) = \Phi(x) = \log Z(x)$ identifies the accessibility entropy with the RSVP scalar field. The field equation $\Delta_g \Phi = V'(\Phi)$ is therefore an equation governing the spatial distribution of accessibility entropy across the manifold. In the projection theory, collapse occurs where the projection π maps regions of high S-variation into a common point in M: the fiber over that point contains admissibility-distinct states with different accessibilities, and the mixing functional Λ reflects that heterogeneity.

19.3 Trajectory Primacy and Semantic Equivalence

A semantic state $x \in X$ carries only instantaneous information: its position in the admissibility manifold and its accessibility field $\mathcal{A}(x)$. Richer semantic content is encoded in trajectories. Define a *semantic trajectory* as a smooth curve $\gamma : [0, T] \to X$ such that $\gamma(t + dt) \in \mathcal{A}(\gamma(t))$ for all t: each step of the trajectory lies within the accessibility field of the current position.

Definition 19.1 (Admissible Trajectory)**.** A smooth curve $\gamma : [0, T] \to X$ is an *admissible trajectory* if $p(\gamma(t+\epsilon) \mid \gamma(t)) > 0$ for all $t \in [0, T)$ and all sufficiently small $\epsilon > 0$. The space of admissible trajectories through x is:

$$\Gamma(x) = \{\gamma : [0, T] \to X \text{ admissible} : \gamma(0) = x\}.$$

The fundamental semantic object is not the state x but the *trajectory equivalence class* $[\gamma]$ of all admissible trajectories through x with the same initial accessibility profile. Two trajectories $\gamma_1, \gamma_2 \in \Gamma(x)$ are semantically equivalent if they have the same first-order accessibility structure: $p(\cdot \mid \gamma_1(t)) = p(\cdot \mid \gamma_2(t))$ for all t in a common initial segment.

This trajectory-first perspective resolves the fiber structure of Part III. The fiber $\pi^{-1}(m)$ over a representational point $m \in M$ is precisely the set of trajectory equivalence classes that project to the same representation. Points in the fiber are not different states; they are different trajectory histories that the projection has identified. The semantic content of each fiber point is the trajectory class that reached it, not the representational value assigned to it. Collapse occurs because projection discards trajectory history, and the discarded history contains admissibility-curvature information that the mixing functional Λ partially recovers.

19.4 Constraint Basins and Semantic Objects

A *semantic object*—a concept, category, or stable meaning unit—is not a point in X or a region of M but an invariant basin of the accessibility field.

Definition 19.2 (Constraint Basin)**.** A measurable subset $B \subseteq X$ is a *constraint basin* if:

$$\mathcal{A}(x) \cap B \neq \varnothing \quad \text{for all } x \in B, \tag{19.3}$$

i.e., every position in B has at least one admissible continuation that remains in B. A constraint basin is *minimal* if it contains no proper invariant sub-basin.

Minimal constraint basins correspond to irreducible semantic units: positions from which the accessibility field returns the system to the basin indefinitely. The filtration $\{X_\lambda\}$ of Chapter 16 organizes basins by admissibility level: X_λ consists of all positions with accessibility entropy above λ, and the connected components of X_λ are the basins accessible at that level.

Semantic collapse at the level of concepts corresponds to basin merging: when projection π identifies points from two distinct basins $B_1, B_2 \subseteq X$ as a common representation $m \in M$, the fiber $\pi^{-1}(m)$ contains trajectories from both basins. The mixing functional Λ measures this inter-basin contamination. A hallucination, in this framework, is a generated continuation that exits the constraint basin of the intended concept while remaining representationally proximate to its projection.

19.5 Semantic Renormalization

The Projection-Collapse Principle is a single-level compression theorem: it relates the admissibility geometry of X to the collapse diagnostics observable in M. Many representational systems apply multiple successive compressions. The natural generalization is a renormalization group structure on the space of admissibility manifolds.

Definition 19.3 (Semantic Renormalization Group)**.** A *semantic renormalization* is a sequence of projections:

$$X \xrightarrow{\pi_1} M_1 \xrightarrow{\pi_2} M_2 \xrightarrow{\pi_3} \cdots \xrightarrow{\pi_n} M_n, \tag{19.4}$$

where each $\pi_k : M_{k-1} \to M_k$ is a smooth surjection (with $M_0 = X$) that integrates out degrees of freedom while preserving the effective admissibility structure relevant at scale k.

At each level k, the projected sectional curvature κ_{π_k} measures the admissibility information destroyed by π_k, and the mixing functional Λ_k measures the observable collapse at that level. The Projection-Collapse Principle applies at each level independently. The composition $\pi_n \circ \cdots \circ \pi_1$ has projected curvature bounded by the sum of the individual curvatures (by the triangle inequality for fiber divergence), giving a total collapse bound:

$$\Lambda_{\text{total}} \leq \sum_{k=1}^{n} \Lambda_k.$$

This is the renormalization group analogue of the data processing inequality: each successive compression can only increase total leakage, never decrease it. The strong form of the Projection-Collapse Principle (Theorem 10.5) applies at each level, making the multilevel collapse computable from the single-level theory.

19.6 The Semantic Action Functional

The preceding structures — accessibility field, accessibility entropy, admissible trajectories, constraint basins, and renormalization — combine into a unified dynamical principle through the semantic action functional.

Definition 19.4 (Semantic Action)**.** For an admissible trajectory $\gamma \in \Gamma(x_0)$ of duration T, define the *semantic action*:

$$\mathcal{S}[\gamma] = \int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t))\, dt, \tag{19.5}$$

where the *semantic Lagrangian* is:

$$\mathcal{L}(x, v) = \frac{1}{2} g_x(v, v) - S(x) = \frac{1}{2} g_{ij}(x)\, v^i v^j - \log \mu(\mathcal{A}(x)). \tag{19.6}$$

The Lagrangian (19.6) has two terms. The kinetic term $\frac{1}{2} g_x(v, v)$ is the squared Fisher speed of the trajectory: it penalizes rapid deformation of the admissibility profile. The potential term $-S(x) = -\log \mu(\mathcal{A}(x))$ is the negative accessibility entropy: it penalizes positions of low admissibility volume, attracting trajectories toward regions of high semantic flexibility. Trajectories minimizing $\mathcal{S}[\gamma]$ subject to fixed endpoints are the solutions to the Euler-Lagrange equation:

$$\ddot{x}^k + \Gamma^k_{ij} \dot{x}^i \dot{x}^j = \frac{1}{2} g^{kl} \partial_l S(x), \tag{19.7}$$

which is the geodesic equation (5.2) with an additional force term driving trajectories toward regions of high accessibility entropy. In the limit $S(x) = \text{const}$, equation (19.7) reduces to the pure geodesic equation (5.2).

The semantic action connects all components of the framework. The Fisher metric supplies the kinetic term and the local geometry. The accessibility entropy supplies the potential and identifies bottlenecks as regions of suppressed volume. Constraint basins are the attractors of the dynamics: trajectories minimizing $\mathcal{S}$ are drawn toward the high-entropy cores of each basin. The projection collapse diagnostics measure the extent to which the projected representation of an action-minimizing trajectory has lost its underlying constraint structure. And the RSVP field equation $\Delta_g \Phi = V'(\Phi)$ emerges as the static limit of the action-minimizing condition when $S = \Phi$ and the trajectory is a constant-speed geodesic: the spatial distribution of accessibility entropy satisfies a field equation that balances diffusion against the potential force.

Chapter 20

Intellectual Foundations and Related Frameworks

The framework developed in this monograph sits at the intersection of several established mathematical and scientific traditions. This chapter positions the theory relative to those traditions, both to acknowledge the work it builds on and to make precise what is new: the geometric theory of admissibility is not a semantic embedding model, a modal logic, a latent-space model, an information bottleneck, or an optimal transport theory. It is a constraint-first field theory from which partial aspects of each of those frameworks emerge as special cases or limiting regimes.

20.1 Information Geometry

The Fisher information metric is the mathematical foundation of the admissibility geometry. The systematic theory of statistical manifolds equipped with the Fisher metric is due to Amari [2, 3], building on the invariance theorem of Chentsov [10]. The α-connections, dual geometries, and the role of the exponential and mixture families in information geometry are developed in full generality in [3].

The present framework uses information geometry in a non-standard way. In standard applications, the statistical manifold parameterizes a family of probability distributions over observed data, and the Fisher metric measures distinguishability of models. Here, the manifold X parameterizes semantic positions, and the distributions $p(\cdot \mid x)$ are over admissible continuations rather than data. The Fisher metric measures semantic distinguishability: how different the accessible futures are at neighboring positions. The invariance theorem (Theorem 4.5) is imported directly from Chentsov's result, re-interpreted in the semantic setting.

20.2 Riemannian Geometry and Submersion Theory

The differential geometry of Riemannian manifolds and smooth maps provides the language for the projection analysis of Part III. Standard references include Lee [18] for an introduction to Riemannian manifolds and O'Neill [22] for semi-Riemannian geometry and the theory of submersions. The O'Neill tensors A and T, the curvature decomposition formulas (C.1)–(C.3), and the relationship between total space, base, and fiber curvature are developed in O'Neill's original paper [21] and the monograph [22].

The Bishop-Gromov comparison theorem (Theorem 6.2) is classical; see for instance Cheeger and Ebin [11] or Petersen [24] for modern treatments. The application to accessibility volume growth in Chapter 6 is new.

20.3 Optimal Transport

The Wasserstein geometry of probability measures is deeply connected to the accessibility structure of the admissibility manifold. Villani [31, 30] develops the optimal transport framework in full generality. The Wasserstein-1 distance appears in the discrete curvature formula (20.2) for Ollivier-Ricci curvature, which is developed in [20].

The accessibility entropy and the geodesic distance d_F have natural optimal transport interpretations: $d_F(x_1, x_2)$ measures the minimum cost of transporting the admissibility distribution at x_1 to that at x_2, in the Fisher-Rao sense. The curvature-divergence inequality (10.2) can be read as a transport cost estimate. The renormalization group structure of Section 17.5 is related to the coarse-graining of transport plans under successive marginalizations.

The present framework is not an optimal transport theory: it does not compute transport plans or solve Monge-Kantorovich problems. But optimal transport provides a natural language for several of its components, and the Wasserstein geometry of $\mathcal{P}(\Omega)$ is the ambient space in which the admissibility manifold is immersed via $\Psi : X \to \mathcal{P}(\Omega)$ (Proposition 3.1).

20.4 Information Bottlenecks and Compression

The information bottleneck method of Tishby, Pereira, and Bialek [29] seeks a compressed representation of an input variable X that preserves maximal information about a target variable Y. The deep variational information bottleneck of Alemi et al. [1] extends this to deep networks via variational inference.

The Projection-Collapse Principle can be interpreted as a geometric generalization of bottleneck theory. In the bottleneck framework, the quality of a compression is measured by the mutual information it preserves. In the present framework, quality is measured by the projected sectional curvature κ_π it destroys: a compression that preserves all mixed curvature is semantically faithful; one that destroys mixed curvature creates observable mixing. The two frameworks agree in spirit — both treat compression as necessarily lossy and ask which losses matter — but differ in their mathematical object: mutual information is a scalar; projected curvature is a field.

20.5 Topological Data Analysis

Persistent homology and its applications to data analysis are developed in Carlsson [9] and the survey of Edelsbrunner and Harer [14]. The admissibility filtration of Chapter 16 and the stability theorem for persistence diagrams (Theorem 16.3) are direct applications of this machinery to the admissibility scalar field Φ.

The connection to sheaf theory, developed briefly in the open problems chapter, is treated in Bredon [7] and Robinson [25]. The cohomological obstruction $H^1(X, \mathcal{A})$ for a sheaf of admissibility structures to glue globally is the formal language for the local-to-global consistency conditions that appear informally throughout the monograph.

20.6 Geometric Deep Learning

Bronstein et al. [8] survey the geometric deep learning framework, which seeks to build neural architectures that respect the symmetries and geometry of their input domains. The reentrant value field architecture of Bohlen [6], discussed in Chapter 13, is an instance of this program applied to coupled symbolic-geometric systems.

The present framework is related but orthogonal: rather than building architectures that respect a given geometry, it studies the geometry that arises from the constraint structure of a semantic domain and asks what any architecture must preserve to be semantically faithful. Geometric deep learning specifies the geometry from outside; admissibility geometry derives the geometry from inside.

20.7 Constraint-Based Accounts of Cognition

Several traditions in cognitive science and theoretical neuroscience treat cognition as navigation through constrained state spaces. The free-energy principle of Friston [15] models perception and action as joint minimization of a variational free energy that bounds surprise. The predictive processing framework of Clark [12] interprets cognition as hierarchical prediction error minimization under generative models.

The admissibility framework is related to these approaches in treating cognition as constraint-governed, but differs in the direction of ontological priority. Free-energy and predictive processing accounts take the generative model as given and derive behavior from it. The admissibility framework takes the constraint structure as primary and derives the generative model from it as the density $p(y \mid x)$ over admissible continuations. The two approaches make different predictions about what happens when representations fail: in free-energy accounts, failure is prediction error; in the admissibility framework, failure is exit from the admissible region, measured geometrically by projected curvature. These are not equivalent descriptions.

20.8 Variational Field Theory

The admissibility action (19.5) and the Euler-Lagrange equations of Chapters 15 and 17 are instances of variational field theory on a Riemannian manifold. Standard treatments include Marsden and Ratiu [19] for the mechanics of constrained systems on manifolds and Bleecker [5] for gauge theory and variational principles.

The RSVP field equations (14.6)–(14.7) are analogous to the coupled scalar-vector field theories of mathematical physics, with the Mexican-hat potential (15.3) playing the role of the Higgs potential. The spontaneous symmetry breaking and vacuum structure of that potential, applied to the admissibility field, produce the bottleneck formation and semantic phase transitions of Chapter 15.

20.9 Process Philosophy and Structural Ontology

The trajectory-first and constraint-first orientation of the framework connects to a long tradition in process philosophy. Whitehead's process and reality [32] treats events and processes as ontologically prior to substances and states; the present framework treats transitions and admissibility as prior to positions and representations. Bergson's creative evolution [4] emphasizes the primacy of duration and continuous becoming over static spatial representation; the admissibility manifold formalizes this intuition as a Riemannian geometry on the space of constraint-governed transitions.

These connections are philosophical rather than technical. The admissibility framework does not import the metaphysics of process philosophy; it shares the structural commitment that the primary objects of a semantic theory should be transitions rather than states, and that meaning should be grounded in the lawfulness of continuation rather than the fidelity of representation.

20.10 Positioning Relative to Existing Frameworks

The admissibility geometry is not any of the following: a semantic embedding model (which represents meaning by position in a vector space and measures semantic similarity by geometric proximity); a modal logic (which models necessity and possibility through accessibility relations between possible worlds, without a metric or curvature structure); a knowledge graph (which represents semantic relations as labeled edges in a discrete network, without a differentiable geometry); a latent-space model (which compresses observations into a low-dimensional representation space, without reference to constraint structure); a free-energy theory (which grounds cognition in the minimization of variational free energy under a generative model); an optimal transport theory (which computes plans for redistributing probability mass at minimum cost); or an information bottleneck framework (which optimizes the tradeoff between compression and task-relevant information).

From each of these, partial aspects can be recovered as special cases or limiting regimes. The Fisher metric reduces to the Euclidean metric in flat regions of the admissibility manifold, recovering the proximity-as-similarity intuition of embedding models. The accessibility field on a discrete graph with Boolean admissibility reduces to a Kripke frame, recovering possible-world accessibility. The renormalization group hierarchy reduces to successive information bottleneck compressions when the admissibility structure is uniform across levels. The semantic action reduces to free-energy minimization in the limit where the accessibility entropy is identified with negative variational free energy.

What the admissibility geometry adds to each of these is the geometric layer: the curvature of the constraint field, the fiber structure of the projection, the mixing functional as a geometric observable, and the reconstruction program that recovers hidden curvature from collapse data. These are not present in any of the existing frameworks, and it is their combination that constitutes the specific contribution of the present work.

20.11 The Ontological Status of Admissibility

Before developing the philosophical consequences of the framework, we address a foundational question that any rigorous reader will pose: what kind of object is the admissibility distribution $p(y \mid x)$? Is it a learned probability model, an empirical frequency, a constraint density, or something more fundamental?

The answer matters because the entire geometric construction rests on p. If p is simply another representation, the theory may be circular. If p is something structurally different from representations, the theory escapes circularity.

Representations and admissibility structures occupy fundamentally different ontological levels. A representation describes states: it assigns labels, coordinates, or vectors to positions in a domain. An admissibility structure describes transformations between states: it specifies which continuations remain lawful from each position. The distinction is not terminological. Formally, let $r : \Omega \to M$ be a representation map assigning labels to states. The admissibility structure is characterized by the family $\{\mathcal{A}(x)\}_{x \in X}$ specifying which continuations are structurally valid from x. While multiple representations may encode the same admissibility structure, the converse does not hold in general: different admissibility structures produce geometrically distinct manifolds (X, g) that no relabeling of states can identify.

This asymmetry is the key. The admissibility field is an invariant under representational change: if r is replaced by any other representation r' that encodes the same continuation constraints, the admissibility geometry is preserved. Representations are coordinate systems imposed upon a deeper field of lawful transitions. The admissibility density $p(y \mid x)$ is not a learned distribution that happens to be useful; it is the encoding of which transitions are structurally valid and with what constraint force. Systems that infer p from data are approximating this structural field from evidence, in the same sense that physical measurements approximate the metric tensor of spacetime from geodesic observations. The approximation may be imperfect; the target it approximates is not itself another representation.

This establishes the ontological ground of the theory. Meaning is not grounded in labels but in the preservation of lawful continuation. The admissibility field generates geometry; representations are secondary projections of that geometry; collapse is the observable signature of a projection that destroys curvature the geometry contains.

20.12 Kinematics and Dynamics of Admissibility

The framework developed in this monograph naturally separates into two complementary layers, and making this separation explicit clarifies the relationship between the geometric and field-theoretic components of the theory.

Definition 20.1 (Semantic Kinematics)**.** *Semantic kinematics* is the study of the geometric structure of the admissibility manifold (X, g) independent of its temporal evolution: the curvature, geodesics, volume growth, fiber structure, and projection geometry that characterize admissibility at a given moment.

Definition 20.2 (Semantic Dynamics)**.** *Semantic dynamics* is the study of how admissibility structures evolve under the RSVP field equations, compression-driven optimization, and constraint flows: the temporal development of Φ, the formation and dissolution of bottlenecks, and the evolution of collapse surfaces.

Parts II and III of this monograph are primarily kinematic. They answer the question: given an admissibility structure, what is its geometry, and what does projection do to that geometry? The RSVP chapters of Part IV are primarily dynamic. They answer the question: how does the admissibility field evolve, and what is the field-theoretic law governing that evolution?

The relationship between these layers mirrors the relationship between geometry and dynamics in physics. The metric tensor of spacetime describes its geometric structure — the curvature of the arena in which trajectories exist. The Einstein field equations describe how matter and energy cause that geometry to evolve. Neither layer is reducible to the other: kinematics tells us the shape of possibility; dynamics tells us how that shape changes.

In the present framework: admissibility geometry provides the static architecture of semantic possibility, encoding which transitions are lawful and at what curvature. RSVP dynamics supplies the equations governing how that architecture evolves under compression, optimization, and learning. The Projection-Collapse Principle is a kinematic theorem: it holds for any admissibility geometry at any moment. The prediction of how collapse evolves during training requires the dynamic layer: the RSVP field equations determine the trajectory of $\Phi(t)$, and through Theorem 15.3, the trajectory of the Fisher metric $g(t)$ and the collapse functional $\Lambda(t)$.

20.13 Semantic Event Horizons

The reconstruction program of Part IV recovers the projected curvature κ_π from collapse measurements Λ. This recovery has a fundamental limit that deserves explicit treatment as a structural feature of the theory rather than a technical inconvenience.

The Jensen-Shannon divergence satisfies $0 \leq D_{\mathrm{JS}}(p,q) \leq \log 2$ for all distributions p and q. The mixing functional Λ, defined using D_{JS}, is therefore bounded above by $\log 2$ regardless of the magnitude of the underlying curvature. As $\kappa_\pi(m)$ grows without bound, $\Lambda(m)$ eventually saturates near its maximum value and ceases to carry curvature information:

$$\lim_{\kappa_\pi(m)\to\infty} \frac{\partial \Lambda(m)}{\partial \kappa_\pi(m)} = 0. \tag{20.1}$$

Beyond the threshold where this saturation sets in, distinct curvature values produce observationally indistinguishable mixing signals. No finite collection of measurements of Λ can distinguish a region of curvature $\kappa_\pi = C$ from one of curvature $\kappa_\pi = 2C$ once both exceed the saturation threshold.

Definition 20.3 (Semantic Event Horizon)**.** A *semantic event horizon* at $m \in M$ is a regime in which $\partial\Lambda(m)/\partial\kappa_\pi(m) \approx 0$: the mixing functional is saturated and further increases in hidden curvature are undetectable from collapse measurements alone.

The existence of semantic event horizons is not a defect of the framework. It is a structural theorem about the limits of any observational program that measures admissibility geometry from the representational side. The bound $\Lambda \leq \log 2$ is tight: there exist configurations where κ_π is large but Λ is at its maximum, and no measurement of Λ alone can determine how far beyond the threshold the curvature has grown.

This establishes a principled ceiling on auditability: geometric features of the admissibility manifold that lie beyond the semantic event horizon cannot be reconstructed from collapse data. They require either direct access to the admissibility field (which may be unavailable), supplementary measurements from a different observable (such as perturbation experiments probing specific tangent directions), or structural assumptions that constrain the geometry beyond what Λ alone can determine.

The horizon is movable. Using alternative divergences with unbounded range—Rényi divergences of order $\alpha > 1$, for instance—the saturation threshold can be pushed to larger values of κ_π, extending the reconstruction regime at the cost of less tractable analysis. The bounded Jensen-Shannon divergence is the natural choice for the core theory because of its metric properties; the extension to unbounded divergences is an open technical program.

20.14 Discrete Admissibility Geometry

The smooth differential geometry developed throughout this monograph applies to semantic systems with continuous parameterizations: probability families that vary smoothly with position, manifolds with well-defined tangent spaces, and projections that are smooth submersions. Many domains of practical significance are not continuous in this sense. Programming languages, formal proof systems, legal codes, and symbolic reasoning architectures have transition structures that are combinatorial: admissibility is determined by syntactic rules, type constraints, or logical entailment, not by a smooth density function.

The continuum framework applies to such systems only as an approximation, valid when the density of admissible transitions is high enough that discrete structure can be idealized as smooth. Where this approximation fails, a discrete version of admissibility geometry is required.

Let $G = (V, E)$ be a directed admissibility graph. Vertices $v \in V$ represent semantic states and directed edges $(v, w) \in E$ represent admissible transitions. The accessibility structure is encoded by the neighborhood:

$$\mathcal{A}(v) = \{w \in V : (v, w) \in E\},$$

with a probability measure $p(\cdot \mid v)$ supported on $\mathcal{A}(v)$ that assigns weights to admissible transitions. The admissibility manifold becomes a weighted directed graph, and the Fisher metric becomes the matrix:

$$g_{ij}(v) = \sum_{w \in \mathcal{A}(v)} \partial_i \log p(w \mid v) \cdot \partial_j \log p(w \mid v) \cdot p(w \mid v),$$

where the derivatives are with respect to parameters of the transition distribution.

Curvature in this discrete setting is captured by Ollivier-Ricci curvature [20], which measures the contraction or expansion of Wasserstein distance under the random walk defined by p. For two adjacent vertices $v_1, v_2 \in V$:

$$\kappa_{\mathrm{OR}}(v_1, v_2) = 1 - \frac{W_1(p(\cdot \mid v_1),\, p(\cdot \mid v_2))}{d(v_1, v_2)}, \tag{20.2}$$

where W_1 is the Wasserstein-1 distance and d is the graph distance. Positive Ollivier-Ricci curvature indicates that the neighborhoods of v_1 and v_2 overlap substantially (the random walks from adjacent vertices spread toward each other); negative curvature indicates that neighborhoods diverge. This is the discrete analogue of the sectional curvature discussion in Chapter 6.

The projection $\pi : X \to M$ becomes a graph homomorphism or, in the sheaf-theoretic generalization, a morphism of sheaves over the admissibility graph. The Projection-Collapse Principle applies in this discrete setting with κ_π replaced by Ollivier-Ricci curvature and Λ replaced by the total variation or Wasserstein mixing between pushforward distributions on the quotient graph.

The smooth and discrete theories share the same conceptual architecture: admissibility structure is primary, representations are projections, curvature measures the information destroyed by projection, and collapse diagnostics recover that curvature from the observable side. The continuum theory emerges as a limit of the discrete theory when the admissibility graph becomes dense and the transition distributions vary smoothly. This limit is analogous to the relationship between discrete lattice field theories and their continuum limits in physics: the discrete version is the more fundamental object, and the smooth manifold is an effective description valid at sufficiently large scales.

20.15 Meaning Without Representation

The central philosophical claim of this monograph is that meaning is not a property of representations. It is an invariant of the admissibility field—a quantity preserved under admissibility-respecting transformations and destroyed by transformations that violate the constraint structure. Representations are projections of that field. They are real, causally efficacious, and practically indispensable. But they are not the primary objects. The field is prior.

This is not a form of anti-representationalism in the philosophical sense. The claim is not that representations do not exist or that they play no role in cognition. The claim is ontological: constraint is prior to representation. The admissibility field generates the geometry from which representations are derived. A representation that faithfully projects the admissibility geometry preserves meaning. One that destroys curvature loses it. The degree of preservation is measured by the Projection-Collapse Principle.

The practical consequence is a design principle: semantic systems should be built constraint-first, not representation-first. The admissibility structure should be specified, approximated, or learned before the representational compression is applied. Adding constraints after the fact, as a downstream repair to a representation that was built without them, is architecturally analogous to adding epicycles: it corrects the specific failure modes that motivated the addition but does not address the underlying geometric cause.

20.16 The Chinese Room Revisited

Searle's Chinese Room argument holds that formal symbol manipulation does not constitute understanding because it lacks intentionality: the symbols are manipulated according to rules but they do not mean anything to the system doing the manipulation [26]. The standard response to this argument within AI research is to add more formal structure: more sophisticated rules, richer representations, more comprehensive coverage of the relevant distinctions.

The admissibility framework suggests a different response. The Chinese Room fails not because it lacks the right symbols but because it operates on representations rather than on the admissibility field those representations project from. The rulebook that the Room operator consults is a description of admissibility relations, not the admissibility relations themselves. A Modal Proofing Kernel adds more pages to the rulebook. It does not move the operator from the representational manifold M back to the admissibility manifold X.

Searle would likely agree with this diagnosis, though not with the remedy proposed here. His own view is that intentionality requires biological substrate. The admissibility framework is agnostic on this question: it does not claim that recovering the admissibility geometry from a compressed representation constitutes understanding in the full philosophical sense. It claims only that the geometric recovery is a necessary condition for the kind of semantic fidelity that logical inference requires. Whether it is sufficient is a deeper question that the present framework leaves open.

20.17 Indexicality and the Admissibility Manifold

Indexical expressions—"I," "here," "now"—are not reducible to descriptions because their semantic content depends on the position of the speaker within the constraint field, not on any property of the token type [23, 16]. The same token "I" expresses different content when uttered from different positions in the admissibility manifold, because the admissible continuations available to different agents differ.

The admissibility manifold provides a natural home for indexical content. The deictic anchor of an indexical is a point $x \in X$: the current position of the utterance within the constraint field. The content of the indexical is a function on X that varies with x: it picks out the agent, location, or time whose position in X is the anchor. The collapse of indexical meaning in representational systems is the projection of this position-dependent function onto a position-independent representation. Any representation that does not encode the anchor point x as part of the semantic state will flatten indexicals into descriptions, producing the conflation of "I" with "the speaker" that is characteristic of indexical collapse.

The fix is not to add a special mechanism for indexical handling. The fix is to include the current position in X as part of the representational state: to ensure that the projection π is equivariant with respect to the group of admissibility-preserving transformations, so that the anchor point is preserved under changes of context.

Chapter 21

Open Problems

21.1 Foundational Problems

The most important open problem is the proof of the Strong Form of the Projection-Collapse Principle. The conjecture states that $\Lambda(m) \asymp \kappa_\pi(m)$ under bounded fiber diameter, Lipschitz continuity of $p(\cdot \mid x)$, and Jensen-Shannon divergence. The weak form is established by the argument of Chapter 6. The strong form requires a matching upper bound: $\Lambda(m) \lesssim \kappa_\pi(m)$. The proof strategy involves a Jensen-Shannon tensorization argument combined with the Lipschitz bound on p and the bounded fiber diameter, but the details require careful treatment of the measure on the fiber and the relationship between within-fiber divergence and representational mixing.

A second foundational problem is the characterization of the admissibility manifolds that arise from natural language, formal reasoning, and perceptual systems. The general construction of X from $p(y \mid x)$ is domain-independent. For specific domains, the geometry of X reflects the specific constraint structure of the domain. Characterizing this geometry—identifying the curvature concentrations, topological invariants, and symmetry groups—is a program of empirical differential geometry applied to semantic systems.

21.2 Geometric Problems

The O'Neill tensor for the submersion $\pi : X \to M$ encodes the mixed curvature in a form amenable to explicit computation [21]. The O'Neill formulas express the sectional curvature of the base M, the fiber $\pi^{-1}(m)$, and the mixed sectional curvature $K_g(v, h)$ in terms of the A-tensor and T-tensor of the submersion. Developing these formulas for the specific case where g is the Fisher metric and π is a neural encoder would give a computationally tractable route to estimating κ_π without relying on triplet measurements.

A related problem is the characterization of the admissibility manifolds arising from specific constraint structures. Boolean constraint satisfaction produces a polytope; linear constraint satisfaction produces a polytope with a natural metric structure from the Fisher information of the feasibility distribution. More complex constraint structures—temporal, modal, causal—produce geometries that have not been systematically characterized.

21.3 Dynamical Problems

The dynamics of the admissibility field under compression-driven optimization is a central open problem. The forward direction is clear: optimization that acts on the representation M rather than the admissibility manifold X flattens the admissibility geometry by collapsing distinctions that are not locally useful. But the precise equations of motion for the admissibility field under gradient descent, contrastive learning, or reinforcement learning have not been derived.

The RSVP framework provides a candidate dynamics, but the connection between the RSVP field equations and the gradient flow on the Fisher metric remains to be made explicit. The question is whether there exists an action functional whose Euler-Lagrange equations are both the RSVP field equations and the natural gradient flow on the space of admissibility densities. If such an action exists, it would unify the constraint-field and optimization-theoretic approaches in a single variational framework.

21.4 Computational Problems

Efficient algorithms for estimating κ_π from finite measurements are needed for practical application of the framework. The current approach—triplet evaluations, entropy measurements, perturbation experiments—provides qualitative evidence about the curvature structure but does not produce quantitative estimates suitable for the tomographic reconstruction of $p(y \mid x)$.

A more systematic approach would apply the tools of topological data analysis, in particular persistent homology, to the collapse graph data. The barcode of a filtration of the collapse graph by collapse rate would provide a multi-scale summary of the curvature structure, identifying topological features of the admissibility manifold at different resolution scales. This approach has been developed for data manifolds in general [9], but its application to admissibility manifolds specifically requires adapting the filtration to the semantics of the domain.

Conclusion: Structure Must Precede Surface

The argument of this monograph can be stated without the mathematical apparatus in a single paragraph. Meaning is the structure of admissible transitions through a hidden constraint field. Representations are projections of that structure. Collapse is what happens when projection destroys curvature the field contains. Collapse diagnostics are measurements of that curvature from the projected side. The inverse problem—recovering the hidden geometry from its observable trace—is the central task of a semantics-aware account of linguistic and cognitive systems.

The mathematical apparatus establishes that this is not a metaphor. The Fisher information metric on the admissibility manifold is positive semi-definite everywhere, well-defined at the locations where Hessian-based approaches fail, and encodes a notion of semantic distinguishability grounded in the structure of admissible transitions. The vertical-horizontal decomposition of the projection identifies the mixed sectional curvature κ_π as the relevant geometric quantity: the curvature that projection destroys and that collapse diagnostics partially recover. The Projection-Collapse Principle establishes the inequality $\Lambda \geq f(\kappa_\pi)$ as a theorem and the equivalence $\Lambda \asymp \kappa_\pi$ as a well-posed conjecture with a clear proof strategy.

The broader lesson is not specific to language models, embedding spaces, or modal logic. It is a lesson about the relationship between constraint, geometry, and representation in any system that encodes structured information in compressed form. Observable failures of representation contain information about hidden constraint structure. That information is recoverable, at least partially, through the systematic analysis of collapse. The geometry can be reconstructed from the trace it leaves when it is destroyed.

This is not a theory about what is missing from current systems. It is a theory about what is present in the collapse patterns those systems already exhibit. The collapse is the message.

Appendix A

Information Geometry: Background and Notation

The Fisher information metric arises in a broader context developed systematically by Amari and colleagues [2]. A statistical manifold is a smooth manifold $\mathcal{S}$ whose points parameterize a family of probability distributions $\{p(\cdot;\theta) : \theta \in \mathcal{S}\}$. The Fisher information metric is the Riemannian metric on $\mathcal{S}$ given by:

$$g_{ij}(\theta) = \mathbb{E}_{x\sim p(\cdot;\theta)}\left[\frac{\partial \log p(x;\theta)}{\partial \theta^i}\frac{\partial \log p(x;\theta)}{\partial \theta^j}\right].$$

In the admissibility manifold of this monograph, the parameter θ is the semantic position $x \in X$ and the family of distributions is the admissibility density family $\{p(\cdot \mid x) : x \in X\}$.

The natural gradient, also due to Amari, is the gradient of a loss function $\ell(\theta)$ preconditioned by the inverse Fisher metric: $\tilde{\nabla}\ell(\theta) = g^{-1}(\theta)\nabla\ell(\theta)$. It is the steepest ascent direction in the metric g, rather than in the Euclidean metric on the parameter space. Learning algorithms that use the natural gradient converge faster in information-geometric terms and are invariant under reparameterization of the statistical manifold.

The α-connections $\nabla^{(\alpha)}$ for $\alpha \in \mathbb{R}$ are a one-parameter family of connections on the statistical manifold, of which the Levi-Civita connection corresponds to $\alpha = 0$ and the exponential and mixture connections correspond to $\alpha = \pm 1$. The ± 1-connections are flat (zero curvature) for exponential family distributions, which is relevant for the analysis of admissibility manifolds arising from constraint structures with exponential family structure.

Appendix B

Differential Geometry: Curvature and Submersions

A smooth submersion $\pi : X \to M$ between Riemannian manifolds is a smooth map whose differential $d\pi_x : T_x X \to T_{\pi(x)} M$ is surjective at every $x \in X$. The vertical bundle $V = \ker(d\pi)$ and its g-orthogonal complement, the horizontal bundle $H = V^{\perp}$, provide the decomposition $TX = V \oplus H$ used throughout this monograph.

The sectional curvature $K_g(\sigma)$ of a Riemannian manifold (X, g) at a point x in the direction of a two-plane $\sigma \subset T_x X$ is defined as:

$$K_g(\sigma) = \frac{g(R(u,v)v, u)}{g(u,u)g(v,v) - g(u,v)^2}$$

for any basis $\{u, v\}$ of σ, where R is the Riemann curvature tensor $R(u,v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w$. The projected sectional curvature $\kappa_\pi(m)$ defined in Chapter 5 integrates $|K_g(v,h)|$ over mixed planes $\sigma = \mathrm{span}\{v, h\}$ with $v \in V_x$ and $h \in H_x$.

Appendix C

The O'Neill Formulas

O'Neill's fundamental equations for a Riemannian submersion relate the curvatures of the total space (X, g), the base (M, g_M), and the fibers to two tensor fields A and T called the O'Neill tensors [21]. Section 5.5 of the main text introduces these tensors in connection with the collapse-driving curvature; this appendix provides the complete formulas and their derivations.

The O'Neill Tensors: Full Definitions

Let $\pi : (X, g) \to (M, g_M)$ be a Riemannian submersion with vertical bundle $V = \ker(d\pi)$ and horizontal bundle $H = V^\perp$. For smooth vector fields E, F on X, write $E = E^H + E^V$ for the decomposition into horizontal and vertical parts. The O'Neill A-tensor and T-tensor are defined as:

$$A_E F = (\nabla_{E^H} F^V)^H + (\nabla_{E^H} F^H)^V,$$
$$T_E F = (\nabla_{E^V} F^H)^V + (\nabla_{E^V} F^V)^H.$$

The A-tensor is skew-symmetric in the sense that $g(A_E F, G) = -g(F, A_E G)$ for horizontal E and arbitrary F, G. The T-tensor at each fiber point is the second fundamental form of the fiber in X.

The Complete Curvature Equations

For orthonormal vectors $h_1, h_2 \in H_x$ (horizontal) and $v_1, v_2 \in V_x$ (vertical), the O'Neill curvature formulas are:

$$K_X(h_1, h_2) = K_M(d\pi(h_1), d\pi(h_2)) - 3\|A_{h_1}h_2\|^2, \tag{C.1}$$

$$K_X(v, h) = -g((\nabla_h A)_v h,\, v) - \|A_h v\|^2 + \|T_v h\|^2, \tag{C.2}$$

$$K_X(v_1, v_2) = K_F(v_1, v_2) - \|T_{v_1}v_2\|^2 + \|T_{v_1}v_1\| \cdot \|T_{v_2}v_2\| + \dots, \tag{C.3}$$

where K_M is the sectional curvature of the base and K_F is the sectional curvature of the fiber. The horizontal-horizontal equation (C.1) shows that the base curvature always exceeds the total-space horizontal curvature by $3\|A_{h_1}h_2\|^2$: the A-tensor reduces curvature as seen from the base.

The mixed equation (C.2) is the key formula for the Projection-Collapse Principle. Rearranging:

$$K_X(v, h) = \|T_v h\|^2 - \|A_h v\|^2 - g((\nabla_h A)_v h,\, v)\,.$$

The collapse-driving term is $-\|A_h v\|^2$: the A-tensor contribution is negative, meaning that non-integrability of the horizontal distribution *reduces* the mixed sectional curvature, coupling the vertical and horizontal directions. The projected sectional curvature (8.2) is therefore controlled by $\|A\|^2$ as established in equation (8.6).

Integrability and Collapse

The horizontal distribution H is integrable (i.e., locally a product $V \times M$) if and only if $A \equiv 0$. In this case all mixed sectional curvatures vanish, $\kappa_\pi = 0$, and the Projection-Collapse Principle gives $\Lambda = 0$: no collapse occurs. This is the geometric condition for a representation to be perfectly faithful on the mixed directions.

When $A \not\equiv 0$, the horizontal distribution is non-integrable: horizontal curves through adjacent fibers do not remain in the same fiber, and the fiber structure twists relative to the base. This twisting is the geometric source of collapse, and $\|A\|^2$ is its quantitative measure.

Appendix D

Proof Details for the Weak Form

We provide a more detailed version of the proof of Theorem 10.4. The key measure-theoretic step is the following lemma.

Lemma D.1. *Let μ be a measure on a product space $\mathcal{Y}_1 \times \mathcal{Y}_2$, and let μ_1, μ_2 be its marginals. If there exists a set $E \subset \mathcal{Y}_1 \times \mathcal{Y}_2$ with $\mu(E) > 0$ on which the density $d\mu/d(\mu_1 \otimes \mu_2)$ differs from 1, then $D(\mu, \mu_1 \otimes \mu_2) > 0$ for any information divergence D that is zero if and only if $\mu = \mu_1 \otimes \mu_2$.*

The proof of the weak form applies this lemma with $\mathcal{Y}_1 \times \mathcal{Y}_2$ corresponding to the product of two admissibility classes in $\pi^{-1}(m)$, and with the mixing in the pushforward distribution providing the set E of positive measure where the density differs from the product. The monotone function f is constructed by tracking the size of E as a function of $\kappa_\pi(m)$ using the continuity of Λ and the compactness of the fiber measure.

Appendix E

Connection to RSVP Field Theory

The Relativistic Scalar-Vector Plenum framework models the constraint structure of a semantic domain as a coupled scalar-vector field over a spacetime manifold $(\mathcal{M}, \eta)$. The scalar field $\Phi : \mathcal{M} \to \mathbb{R}$ is the admissibility density; the vector field $V : \mathcal{M} \to T\mathcal{M}$ is the accessibility flow. The RSVP field equations are:

$$\Box\Phi = \lambda(\Phi^2 - v^2)\Phi - \nabla_\mu V^\mu, \tag{E.1}$$

$$\nabla_\nu F^{\mu\nu} = J^\mu[\Phi], \tag{E.2}$$

where $\Box$ is the wave operator on $(\mathcal{M}, \eta)$, $F^{\mu\nu}$ is the field strength of V, and $J^\mu[\Phi]$ is the current sourced by the scalar field. These equations enforce a Mexican-hat potential for Φ, producing spontaneous symmetry breaking at locations where the admissibility density falls below the threshold v: the semantic bottlenecks discussed in Chapter 3.

The connection to the Fisher metric is through the identification $g_{ij}(x) = \partial_i\partial_j\Phi(x)$ near local maxima of Φ, where the scalar field is strictly concave and the Hessian is negative definite. Near maxima, the RSVP scalar field approximates the admissibility density, and its Hessian approximates the Fisher metric. Away from maxima, the Fisher metric is the correct object. The RSVP framework provides the dynamics; the Fisher metric provides the static geometry; they agree in the high-admissibility regions where the system spends most of its time.

Bibliography

[1] A. Alemi, I. Fischer, J. Dillon, and K. Murphy, "Deep variational information bottleneck," *International Conference on Learning Representations*, 2017.

[2] S. Amari, *Information Geometry and Its Applications*, Springer, Tokyo, 2016.

[3] S. Amari and H. Nagaoka, *Methods of Information Geometry*, American Mathematical Society, Providence, 2000.

[4] H. Bergson, *Creative Evolution*, translated by A. Mitchell, Henry Holt, New York, 1911.

[5] D. Bleecker, *Gauge Theory and Variational Principles*, Addison-Wesley, Reading MA, 1981.

[6] K. Bohlen, "Reentrant value fields as delayed coupled reaction-diffusion systems on finite graphs," arXiv:2605.03940v2, 2026.

[7] G. E. Bredon, *Sheaf Theory*, second edition, Springer, New York, 1997.

[8] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: grids, groups, graphs, geodesics, and gauges," arXiv:2104.13478, 2021.

[9] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society* 46 (2009), 255–308.

[10] N. N. Chentsov, *Statistical Decision Rules and Optimal Inference*, American Mathematical Society, Providence, 1982.

[11] J. Cheeger and D. Ebin, *Comparison Theorems in Riemannian Geometry*, North-Holland, Amsterdam, 1975.

[12] A. Clark, *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford University Press, New York, 2016.

[13] S. Denham, "Semantic collapse in embedding space: neighborhood semantics and entropic drift," 2026.

[14] H. Edelsbrunner and J. Harer, *Computational Topology: An Introduction*, American Mathematical Society, Providence, 2010.

[15] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience* 11 (2010), 127–138.

[16] D. Kaplan, "Demonstratives," in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*, Oxford University Press, Oxford, 1989, pp. 481–563.

[17] S. A. Kripke, "Semantical analysis of modal logic I: normal modal propositional calculi," *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9 (1963), 67–96.

[18] J. M. Lee, *Introduction to Riemannian Manifolds*, second edition, Springer, Cham, 2018.

[19] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, second edition, Springer, New York, 1999.

[20] Y. Ollivier, "Ricci curvature of Markov chains on metric spaces," *Journal of Functional Analysis* 256 (2009), 810–864.

[21] B. O'Neill, "The fundamental equations of a submersion," *Michigan Mathematical Journal* 13 (1966), 459–469.

[22] B. O'Neill, *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press, New York, 1983.

[23] J. Perry, *The Problem of the Essential Indexical and Other Essays*, Oxford University Press, New York, 1993.

[24] P. Petersen, *Riemannian Geometry*, third edition, Springer, Cham, 2016.

[25] M. Robinson, *Topological Signal Processing*, Springer, Heidelberg, 2014.

[26] J. R. Searle, "Minds, brains, and programs," *Behavioral and Brain Sciences* 3 (1980), 417–457.

[27] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal* 27 (1948), 379–423.

[28] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston, Washington DC, 1977.

[29] N. Tishby, F. Pereira, and W. Bialek, "The information bottleneck method," arXiv:physics/0004057, 2000.

[30] C. Villani, *Topics in Optimal Transportation*, American Mathematical Society, Providence, 2003.

[31] C. Villani, *Optimal Transport: Old and New*, Springer, Berlin, 2009.

[32] A. N. Whitehead, *Process and Reality*, Macmillan, New York, 1929.