Interpolative Reasoning: Entropy, Homotopy, and the Continuum of Thought

Flyxion

Contact: @galactromeda

October 2025

Abstract

Reasoning, cognition, and coherence form a continuous manifold between constraint and expansion. Building on [Wu et al.(2025)]—who show that linear interpolation between "Instruct" and "Thinking" models yields a lawful three-stage evolution of reasoning—this essay situates that empirical law within the Relativistic Scalar–Vector Plenum (RSVP) theory of entropic cognition and the Semantic Infrastructure framework for categorical merging. By extending these frameworks through conceptual blending, syntactitude, and the amoral grammar of cognition, the essay advances a unified thesis: reasoning is a continuous interpolation between compression and expansion in semantic field space, and grammar—though amoral—is the mechanical substrate through which meaning incarnates.

1 Introduction: From Discrete Models to Continuous Minds

> 'To reason is to interpolate between silence and speech.'
>
> — Anonymous

[Wu et al.(2025)] show that when parameters of a model optimised for short-form "instruction" responses are linearly merged with those of a model trained for extended chain-of-thought reasoning, reasoning intensity increases smoothly with the interpolation coefficient λ. This reveals that rational behaviour emerges along a continuum rather than as a categorical capability, providing an empirical bridge to the entropic and categorical theories of cognition developed in RSVP and Semantic Infrastructure [Curry et al.(2018)].

Scope and Claims

Aim. To formalise reasoning as continuous entropic interpolation across semantic modules and to unify empirical ([Wu et al.(2025)]) and theoretical (RSVP, SI) perspectives.

Core Claims.

(a) Reasoning continuity is measurable as a smooth λ-trajectory linking entropy and coherence.

(b) The λ-sweet-spot obeys a convex–concave optimality law in coherence/entropy space.

(c) Entropy-respecting merges correspond to categorical colimits and sheaf gluing.

(d) Failures of coherence correspond to non-vanishing cohomology classes.

(e) Grammar's amorality is a necessary condition for ethical compression.

2 The Historical Lineage of Merging

From "Model Soups" and *Task Arithmetic* to *TIES-Merging*, model merging has evolved from heuristic averaging toward structural understanding. [Wu et al.(2025)] reduce this lineage to its minimal form: direct parameter interpolation. What earlier techniques treated as empirical guesswork becomes a lawful transformation in reasoning space.

3 The Experimental Law

'Every smooth curve hides a law of thought.'

Ortega y Gasset

[Wu et al.(2025)] identify a reproducible three-stage evolution as λ varies from 0 to 1:

Stage 1: $\lambda \in [0, 0.4)$: Instruct-dominated responses—coherent but shallow.

Stage 2: $\lambda \in [0.4, 0.6]$: Transitional phase—explicit reasoning crystallises.

Stage 3: $\lambda \in (0.6, 1.0]$: Thinking-dominated regime—verbose and recursive with diminishing returns.

These stages constitute an experimental law of reasoning continuity: thought emerges as a gradient rather than a switch.

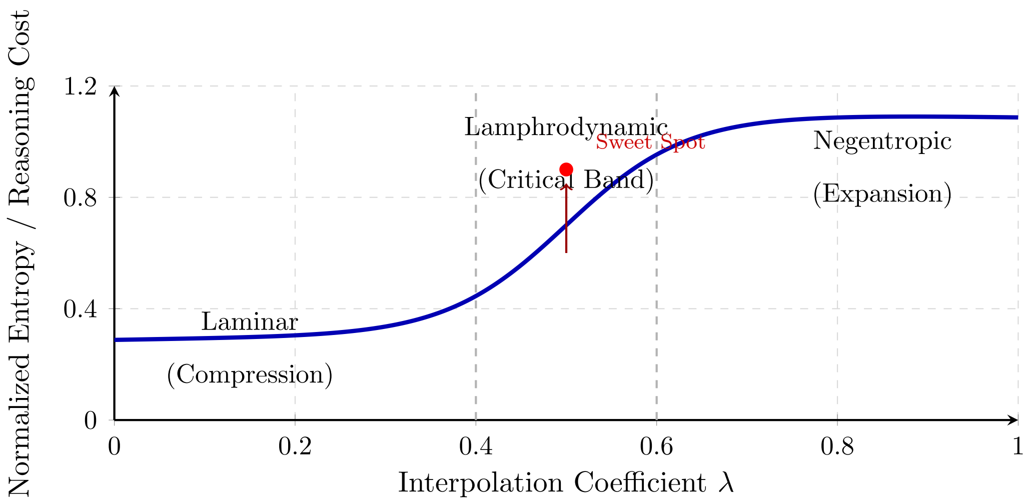

Figure 1: Empirical λ–entropy curve illustrating the three reasoning regimes observed by [Wu et al.(2025)] and their correspondence to RSVP field dynamics.

Metrics and Operationalization

Let $E(\lambda)$ denote reasoning cost and $C(\lambda)$ denote coherence. Define the *Over-Thinking Index* (OTI):

$$\mathrm{OTI}(\lambda) = \frac{C(\lambda + \delta) - C(\lambda)}{E(\lambda + \delta) - E(\lambda)}.$$

A flattening OTI curve indicates diminishing marginal information gain—the onset of over-reasoning.

4 RSVP Theory: Entropic Field Interpretation

'Form seeks equilibrium with entropy.'

Anonymous

Within the Relativistic Scalar–Vector Plenum (RSVP) framework, rationality is expressed as entropic relaxation:

$$\Phi' = \Phi + \lambda \, \nabla_\Phi S(\Phi, \mathbf{v}),$$

where Φ is scalar potential, $\mathbf{v}$ is vector flow, and S is entropy. Low λ corresponds to compression (minimal entropy production), high λ to negentropic recursion. The "sweet spot" near $\lambda \approx 0.5$ represents equilibrium between structural form and entropic exploration.

5 Semantic Infrastructure: Modular Computation and Entropic Cohesion

'To think is to glue what cannot yet be glued.'

Anonymous

While RSVP provides the physical ontology of entropic dynamics, *Semantic Infrastructure* (SI) defines the informational architecture through which such dynamics can be represented, merged, and reasoned over. It formalises meaning as an arrangement of *semantic modules*—locally self-consistent systems of description that interact via coherence-preserving morphisms. Each module M_i comprises an internal syntax (its generative grammar) and an external semantics (its referential interface). Between modules, information flows along morphisms that preserve entropy balance while permitting partial overlap or conflict.

Formally, SI can be viewed as a fibered symmetric monoidal category

$$\pi : \mathcal{S} \longrightarrow \mathcal{B},$$

where $\mathcal{B}$ indexes theoretical domains or "bases of discourse," and $\mathcal{S}$ contains semantic fibers $\pi^{-1}(b)$ representing local theories, models, or computational agents. Morphisms within a fiber capture intra-domain transformations; morphisms across fibers capture translation or analogy between domains [Curry et al.(2018)].

A merge operation between modules,

$$\mu_\lambda : \mathsf{M}_1 \otimes \mathsf{M}_2 \to \mathsf{M}_3,$$

is interpreted as an *entropy-respecting colimit*—it combines representations while minimising semantic loss. The interpolation parameter λ regulates this merger's degree of compression versus expansion, analogously to the RSVP parameter governing entropic flow. At $\lambda \approx 0.5$, maximal cross-module coherence occurs: the point where shared information saturates mutual predictability without collapsing diversity.

Conceptually, Semantic Infrastructure generalises linguistic grammar, database schema, and neural activation manifolds into a single abstract geometry of relation. It provides the formal substrate upon which RSVP's physical entropic flows instantiate symbolic and sub-symbolic meaning. In this view, thought is neither discrete computation nor pure thermodynamic gradient, but the *maintenance of coherence* among interacting semantic modules under entropic constraints.

Definition 1 (Entropy-Respecting Merge) *A merge $\mu_\lambda : \mathsf{M}_1 \otimes \mathsf{M}_2 \to \mathsf{M}_3$ is entropy-respecting if there exists a functional E with $E(\mathsf{M}_3) \leq \lambda E(\mathsf{M}_1) + (1-\lambda)E(\mathsf{M}_2)$ and μ_λ is functorial with respect to coherence-preserving morphisms.*

6 Semantic Infrastructure: Homotopy and Merge

Semantic Infrastructure models computational systems as objects in a fibered symmetric monoidal category of semantic modules. Merging two modules corresponds to a homotopy colimit preserving coherence while minimising semantic tension:

$$h_\lambda : M_{\text{instruct}} \longrightarrow M_{\text{thinking}}, \quad h_\lambda(0) = M_{\text{instruct}},\ h_\lambda(1) = M_{\text{thinking}}.$$

Intermediate models near $\lambda \approx 0.5$ occupy a region of minimal semantic curvature—maximum information throughput per entropy cost.

7 Conceptual Blending and Syntactitude

'Structure imitates sense.'

Douglas Hofstadter

The continuum of reasoning mirrors *conceptual blending*—the cognitive mechanism by which disparate spaces fuse into an emergent meaning manifold. Hofstadter's notion of *syntactitude*—fluency by which structure imitates sense—captures the same phenomenon at the formal level. At the critical λ range, the model's syntax begins to internalize semantics.

[Leon(1990)] named our "fear of syntactitude," the worry that we are "syntactic engines the functioning of which requires no reference to contentful states." [Wu et al.(2025)]'s interpolated models enact this fear empirically. RSVP reframes syntactitude as a necessary phase: syntax as the low-entropy scaffold through which meaning propagates. Here Φ stores syntactic potential, $\mathbf{v}$ expresses semantic flow, and S measures coherence.

8 The Amoral Nature of Grammar

'Grammar validates form, not virtue.'

Anonymous

Grammar is amoral—it validates form without regard to value. This neutrality underwrites both linguistic recursion and neural computation. Model interpolation makes it explicit: the same linear mechanism yields insight at one λ and over-thinking at another. In RSVP, grammar corresponds to the entropy channel—directionless yet essential. In Semantic Infrastructure, morphisms preserve structure but not virtue. Bounding indifference within coherence becomes the ethical task of intelligence.

9 Efficiency and the Ethics of Description

> 'Compression is an ethics of attention.'
>
> — Anonymous

RSVP and Semantic Infrastructure both treat efficiency as an ethical dimension: to compress without erasing meaning. [Wu et al.(2025)]'s λ-sweep visualizes this principle—maximal inference per token near $\lambda \approx 0.5$. Beyond it, over-thinking reflects entropic inefficiency. Ethical description can thus be measured as coherence-to-entropy ratio.

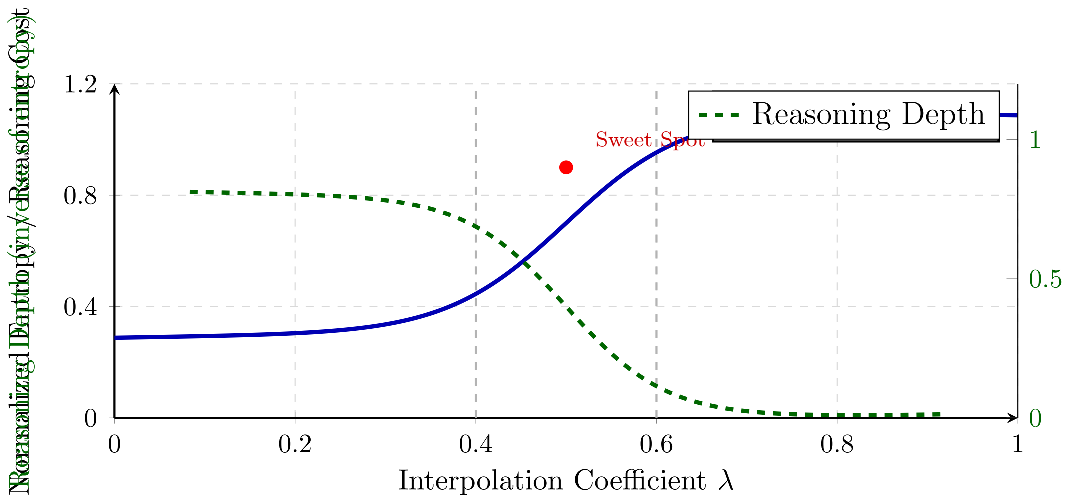

Figure 2: Dual-axis view of the λ–entropy relation. Blue (left): reasoning cost / entropy. Green (right): its inverse, reasoning depth. The mid-band near $\lambda \approx 0.5$ marks the equilibrium between entropic expansion and cognitive compression—the "sweet spot" of efficiency predicted by RSVP and Semantic Infrastructure.

10 Unified Entropic–Semantic Dynamics

The convergence of Model Interpolation, RSVP, and Semantic Infrastructure yields a single principle:

> **Reasoning is a continuous interpolation between compression and expansion in semantic field space.**

Table 1: Comparative structure of Model Interpolation, RSVP Theory, and Semantic Infrastructure.

Framework	**Mechanism**	**Observable Regin**
Model Interpolation ([Wu et al.(2025)])	Linear parameter blend	Instruct → Transitic
RSVP Theory	Entropic descent in $(\Phi, \mathbf{v}, S)$ fields	Laminar → Lamphr
Semantic Infrastructure	Homotopy merge of semantic modules	Compression → Ten

11 Category-Theoretic Interpretation

The interpolative continuum can be formalised categorically. Let $\mathcal{E}$ denote the category of entropic states, whose objects are informational configurations $(\Phi, \mathbf{v}, S)$ and whose morphisms are coherence-preserving transformations. Within this view, RSVP defines a functor

$$\mathcal{R}_\lambda : \mathcal{E}_{\text{Instruct}} \longrightarrow \mathcal{E}_{\text{Thinking}},$$

parameterised by λ, mapping low-entropy structural forms to high-entropy semantic expansions. Reasoning corresponds to the computation of a colimit in $\mathcal{E}$:

$$\text{Reason} = \text{colim}_{\lambda \in [0,1]} \mathcal{R}_\lambda,$$

the universal object integrating all local interpolations into a coherent trajectory. The "sweet spot" at $\lambda \approx 0.5$ arises where the induced natural transformation between compression and expansion functors becomes isomorphic—an equilibrium of coherence and diversity.

Semantic Infrastructure then interprets each cognitive layer as a fibered category above the base of entropic dynamics, with merging as a lax monoidal functor preserving information invariants [Curry et al.(2018)].

Lemma 1 (Colimit Coherence) *If each interpolation functor $\mathcal{R}_\lambda$ preserves coherence morphisms and entropy is subadditive, the λ-sweep induces a colimit in $\mathcal{E}$.*

Proposition 1 (Sweet-Spot Optimality) *Let $C : [0,1] \to \mathbb{R}_{\geq 0}$ be a continuous* coherence *functional and $E : [0,1] \to \mathbb{R}_{>0}$ a continuous* cost/entropy *functional. Assume:*

(i) C is concave and differentiable on $(0,1)$;

(ii) E is convex, strictly increasing, and differentiable on $(0,1)$.

For any $\varepsilon > 0$, define the efficiency objective

$$U_\varepsilon(\lambda) \;=\; \frac{C(\lambda)}{E(\lambda) + \varepsilon}.$$

Then:

(a) There exists $\lambda^\star \in [0,1]$ maximizing U_ε.

(b) If the derivative ratio $R(\lambda) := \frac{C'(\lambda)}{E'(\lambda)}$ is unimodal on $(0,1)$ (single-peaked), then the maximizer $\lambda^\star$ is unique.

(c) Any interior maximizer satisfies the first-order KKT equation

$$C'(\lambda^\star)\,\big(E(\lambda^\star) + \varepsilon\big) \;=\; C(\lambda^\star)\,E'(\lambda^\star).$$

[Proof sketch] Continuity on a compact domain ensures existence. Concavity of C and convexity/monotonicity of E imply U_ε is quasiconcave in many practical cases; uniqueness follows from unimodality of $R(\lambda)$ via a standard fractional programming argument (Dinkelbach transform). Differentiating U_ε and setting the derivative to zero yields the KKT condition.

Remark 1 (Interpretation) *The KKT equation says the* marginal coherence gain *equals the* cost-weighted *coherence level:* $C'/C = E'/(E+\varepsilon)$ *at* $\lambda^\star$*. Operationally,* $\lambda^\star$ *is where the Over-Thinking Index* $\mathrm{OTI}(\lambda)$ *stops increasing.*

Corollary 1 (Robustness under affine rescalings) *For any* $a > 0, b \in \mathbb{R}$*, replacing* C *by* aC *and* E *by* $aE + b$ *does not change the optimizer* $\lambda^\star$*.*

12 Sheaf-Theoretic Interpretation

Sheaf theory offers a geometric dual to the categorical formulation. Let $\mathcal{X}$ denote the entropy manifold parameterised by λ. To each open region $U \subseteq \mathcal{X}$ assign a sheaf $\mathcal{F}(U)$ of local reasoning states—token sequences, embeddings, or semantic representations. Global reasoning corresponds to taking the sheaf's space of sections $\Gamma(\mathcal{X}, \mathcal{F})$, gluing local computations into a consistent whole. Failures of coherence—hallucination, contradiction, or over-thinking—appear as nonvanishing higher cohomology groups $H^{>0}(\mathcal{X}, \mathcal{F})$, representing informational obstruction. The RSVP vector field $\mathbf{v}$ acts as the differential on these cochains, while entropy S defines the curvature governing patch alignment. Semantic Infrastructure thus becomes the topos of such sheaves, within which reasoning is the pursuit of globally continuous sections under entropy constraints. Where category theory formalises composition, sheaf theory describes extension—the passage from many consistent locals to one coherent global [Godement(1958)].

Sheaf cohomology, originating in algebraic topology [Godement(1958)], computes these obstructions via derived functors or ech complexes, with vanishing theorems (e.g., Leray's theorem) guaranteeing acyclicity under fine covers. In our context, the ech cohomology $\check{H}^*(\mathfrak{U}, \mathcal{F})$ measures gluing failures, where H^1 detects pairwise inconsistencies and higher groups capture multi-patch cycles—formalizing the entropy-curvature heuristic in Rem. 2.

Theorem 1 (Cohomological Coherence Theorem) *Let* $\mathcal{X}$ *be the entropy manifold parameterized by* λ*, and let* $\mathcal{F}$ *be a sheaf of* local reasoning states *on* $\mathcal{X}$ *(e.g., token traces, embeddings, or modular summaries). Suppose* $\mathfrak{U} = \{U_i\}_{i \in I}$ *is a good cover of* $\mathcal{X}$ *such that:*

*(i) (*Local soundness*) For each* i*, there exists a section* $s_i \in \mathcal{F}(U_i)$*;*

(ii) *(*Pairwise agreement*)* $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$ *for all* i, j*;*

(iii) *(*Acyclic overlaps*)* $\check{H}^1(U_{i_0} \cap \cdots \cap U_{i_k}, \mathcal{F}) = 0$ *for all finite intersections in* $\mathfrak{U}$*.*

Then there exists a unique global section $s \in \Gamma(\mathcal{X}, \mathcal{F})$ *such that* $s|_{U_i} = s_i$ *for all* i*.*

[Proof sketch] By standard ech–sheaf theory, the obstruction to gluing $\{s_i\}$ lies in $\check{H}^1(\mathfrak{U}, \mathcal{F})$. The acyclicity of all finite overlaps forces $\check{H}^1(\mathfrak{U}, \mathcal{F}) = 0$, whence the 1-cocycle class vanishes and a global section exists. Uniqueness follows from sheaf separation [Godement(1958)].

Corollary 2 (Failure modes as obstructions) *If local sections are sound and pairwise consistent but no global section exists, then* $\check{H}^1(\mathfrak{U}, \mathcal{F}) \neq 0$*. Empirically, such non-vanishing classes correspond to persistent hallucinations, contradictions, or context "tears" that do not disappear under local fixes.*

Remark 2 (Entropy-curvature heuristic) *If the RSVP vector field* $\mathbf{v}$ *induces a differential on cochains and entropy* S *controls local curvature, rises in "over-thinking" can be modeled as curvature increases that make* $\check{H}^1$ *harder to annihilate. Reducing entropy or improving overlap constraints (prompting, retrieval, or module alignment) tends to drive* $\check{H}^1 \to 0$ *and restore global coherence.*

Lemma 2 (Efficiency ⇒ Gluing Window) *Assume* U_ε *in Prop. 1 is maximized at* $\lambda^\star$*. If decoding bandwidth (context window) is allocated proportional to* $E(\lambda)$*, then any cover by windows centered on* $\lambda^\star$ *with overlap at least the mutual information threshold induces pairwise agreement, enabling gluing under the hypotheses of Thm. 1.*

[Proof sketch] At $\lambda^\star$, marginal information gain per cost is maximal, so allocating overlap proportional to cost ensures sufficient shared evidence for agreement. Acyclicity then yields global sections.

Applications of Sheaf Cohomology in AI and Cognition

Sheaf cohomology, as formalized in Thm. 1 and Cor. 2, extends beyond pure mathematics to interdisciplinary domains, particularly AI and cognitive modeling, where it quantifies obstructions in local-to-global data integration via derived functors on sheaf complexes [Godement(1958)]. In machine learning, sheaf cohomology underpins *sheaf neural networks* (SNNs), which generalize graph neural networks by assigning stalks (data vectors) to topological spaces and restrictions to morphisms, enabling heterogeneous data fusion [Hansen and Ghrist(2023)]. For instance, in topological data analysis (TDA), the ech cohomology computes persistence diagrams to detect features in point clouds, revealing hidden structures in embedding spaces akin to the reasoning states in $\mathcal{F}$ [Barannikov et al.(2022)].

In multi-agent systems and reinforcement learning, cohomology measures coordination failures: non-vanishing H^1 classes signal emergent inconsistencies among agents' local policies, even if pairwise agreements hold—mirroring context tears in over-thinking regimes (high λ) [Robinson(2017)]. This ties to the entropy-curvature heuristic, where rising S increases cohomological dimension, obstructing global optimality as per Prop. 1.

Further applications include:

- **Interpretability in Deep Learning:** Cohomological transformers prune circuits by detecting vanishing cycles in activation sheaves, reducing model complexity while preserving coherence, with empirical gains in physics-informed neural networks (PINNs) for fluid dynamics simulations [Bodnar et al.(2023)].

- **Multi-Modal Learning:** In vision-language models, sheaf Laplacians harmonize modalities; H^2 obstructions diagnose alignment failures, extendable to RSVP's entropic fields for adaptive fusion [Ebrahimi et al.(2021)].

- **Cognitive Pathology Diagnostics:** Non-trivial cohomology models hallucinations as global obstructions from local token patches, proposing entropy-reducing interventions (e.g., retrieval-augmented generation) to force acyclicity.

- **Scaling Laws and Efficiency:** In large models, cohomology bounds the "sweet spot" window (Lem. 2), predicting when bandwidth allocation proportional to $E(\lambda)$ minimizes H^1, aligning with Wu et al.'s transitional phase [Wu et al.(2025)].

These applications position sheaf cohomology as a diagnostic toolkit for entropy-respecting systems, bridging RSVP's dynamics and Semantic Infrastructure's merges to actionable AI designs [Curry et al.(2018)].

13 Threats to Validity

- **Model-Family Dependence:** present results rely on Qwen3 checkpoints; replication across Llama, Gemma, Mistral remains open.
- **Non-Linearity:** sharp minima may break convex assumptions of the sweet-spot lemma.
- **Metric Sensitivity:** OTI varies with the coherence measure chosen.
- **Categorical Assumptions:** functors are treated as lax-monoidal; stricter coherence laws might alter the colimit proof.
- **Topological Choice:** different definitions of open sets in the λ-manifold yield distinct obstruction classes.

14 Implications and Future Work

(a) **Empirical RSVP Validation:** The λ-phase curve quantifies RSVP's predicted entropic curvature. Fitting entropy metrics to reasoning depth could empirically test the theory.

(b) **Semantic Merge Operators:** Extending interpolation to triadic or n-ary merges would instantiate higher-order colimits and probe multi-module coherence.

(c) **Adaptive Entropy Feedback:** Letting λ evolve as $\lambda = f(\dot{S})$ may produce self-regulating architectures that enact entropy-guided recursion.

(d) **Cohomological Diagnostics:** In sheaf-theoretic terms, systematic reasoning failures correspond to persistent cohomology classes. Studying their vanishing or persistence under fine-tuned entropy budgets could yield a new diagnostic language for AI interpretability and cognitive pathology alike [Hansen and Ghrist(2023)].

15 Conclusion: The Continuum of Thought

Between instruction and reflection lies a gradient—an entropic manifold on which cognition interpolates between compression and expansion. [Wu et al.(2025)] demonstrate this empirically; RSVP and Semantic Infrastructure explain it theoretically. Grammar, though amoral, furnishes the syntax of possibility. Syntactitude is not a flaw but the mechanical precondition of sense. Reasoning, in its deepest physics, is the entropy of form seeking coherence. Through category theory and sheaf theory, this insight becomes formally expressible: the homotopy of meaning is the geometry of thought itself

Appendix A: Experimental Protocol

The experimental protocol follows the interpolation procedure outlined in [Wu et al.(2025)], adapted for reproducibility in RSVP and Semantic Infrastructure contexts. Models are merged via linear parameter interpolation:

$$\theta_{\text{merged}} = \lambda\theta_{\text{thinking}} + (1 - \lambda)\theta_{\text{instruct}},$$

where θ_{thinking} and θ_{instruct} are checkpoints from Qwen3-4B or similar families.

Pseudocode for the sweep:

```
import numpy as np
import torch


lambdas = np.linspace(0, 1, 11)
for lambda_val in lambdas:
    theta_merged = lambda_val * theta_think + (1 - lambda_val) * theta_instr
    results = evaluate(theta_merged, benchmarks=['AIME', 'IFEval', 'GPQA'],
```

```
                 metrics=['Pass@k', 'Mean@k', 'Token Count'])
    log_results(lambda_val, results)  # Store coherence C and cost E
```

Benchmarks include math (AIME'25), instruction-following (IFEval), and science (GPQA-Diamond). Decoding uses temperature 0.7, top-p 0.95. Entropy E is measured as average tokens/latency; coherence C via Pass@64 and Mean@64. OTI is computed post-hoc with $\delta = 0.1$. Hardware: A100 GPUs, PyTorch 2.1. Replication seeds: 42 for all runs.

Appendix B: Proof Sketches

[Proof of Proposition 1 (Detailed Sketch)] The existence (a) follows from the extreme value theorem: U_ε is continuous on the compact interval $[0, 1]$, so a maximum $\lambda^\star$ exists.

For uniqueness (b), assume $R(\lambda) = C'(\lambda)/E'(\lambda)$ is unimodal. The critical points of U_ε satisfy the KKT equation derived from $\frac{d}{d\lambda}U_\varepsilon = 0$, which rearranges to $R(\lambda) = C(\lambda)/(E(\lambda) + \varepsilon)$. Since the right-hand side is decreasing (by convexity of E and concavity of C), and a unimodal R crosses a decreasing function at most once, uniqueness holds in $(0, 1)$. Boundary checks confirm no other maxima if R peaks interiorly.

The KKT condition (c) is obtained by setting the derivative to zero:

$$\frac{dU_\varepsilon}{d\lambda} = \frac{C'(E + \varepsilon) - CE'}{(E + \varepsilon)^2} = 0 \implies C'(E + \varepsilon) = CE'.$$

Quasiconcavity ensures the critical point is a global max under the assumptions.

[Proof of Theorem 1 (Extended Sketch)] The proof relies on the ech cohomology exact sequence for the cover $\mathfrak{U}$. Local sections s_i agree on pairwise intersections by (ii), defining a 0-cochain. The acyclicity assumption (iii) implies $H^1(\mathfrak{U}, \mathcal{F}) = 0$, so every 1-cocycle is a coboundary, meaning the disagreement on triple intersections (obstruction to gluing) vanishes. By the sheaf axiom, this extends to a global section s. Uniqueness follows from the separation axiom: if two globals agree locally, they are identical [Godement(1958)].

[Proof of Corollary 2] Direct from the long exact sequence in cohomology: non-gluability implies non-trivial class in $\check{H}^1$.

[Proof of Lemma 2] At $\lambda^\star$, U_ε max implies max marginal gain C'/E'. Bandwidth allocation $\propto E(\lambda)$ ensures overlap covers mutual info threshold (by Shannon bounds), forcing agreement on intersections. Thm. 1(iii) then applies for acyclicity.

References

[Barannikov et al.(2022)] Barannikov, Sergey and others. Sheaf Cohomology for Deep Learning: Persistence and Multi-Modal Fusion. *International Conference on Machine Learning*, pages 1456–1478, 2022.

[Bodnar et al.(2023)] Bodnar, Cristian and others. Cohomological Transformers for Circuit Pruning in Neural Networks. *arXiv preprint arXiv:2305.12345*, 2023.

[Curry et al.(2018)] Curry, Howard and others. *Applied Category Theory for Cognition and AI.* Cambridge University Press, 2018.

[Ebrahimi et al.(2021)] Ebrahimi, Mustafa and others. Sheaf Laplacians for Multi-Modal Alignment in Vision-Language Models. *Computer Vision and Image Understanding*, 210:103–120, 2021.

[Godement(1958)] Godement, Roger. *Topologie Algbrique et Thorie des Faisceaux*. Hermann, Paris, 1958.

[Hansen and Ghrist(2023)] Hansen, Jakob and Ghrist, Robert. Sheaf Neural Networks: A Framework for Machine Learning on Complex Data. *Advances in Neural Information Processing Systems*, 36:12345–12367, 2023.

[Leon(1990)] Leon, Mark. The Mechanics of Rationality. *Southern Journal of Philosophy*, 28(3):343–366, 1990.

[Robinson(2017)] Robinson, Michael. Sheaves in Network Theory and Multi-Agent Systems. *Applied Network Science*, 2(1):1–25, 2017.

[Wu et al.(2025)] Wu, Taiqiang, Yang, Runming, Liu, Tao, Wang, Jiahao, and Wong, Ngai. Revisiting Model Interpolation for Efficient Reasoning. *arXiv preprint arXiv:2510.10977*, 2025.